Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment

Yannick Thiery, Jean-Philippe Malet, Simone Sterlacchini, Anne Puissant, Olivier Maquaire

▶ **To cite this version:**

Yannick Thiery, Jean-Philippe Malet, Simone Sterlacchini, Anne Puissant, Olivier Maquaire. Landslide susceptibility assessment by bivariate methods at large scales: Application to a complex mountainous environment. Geomorphology, 2007, 92 (1-2), pp.38-59. 10.1016/j.geomorph.2007.02.020 . hal-00276804

HAL Id: hal-00276804
https://hal.science/hal-00276804

Submitted on 2 May 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Elsevier Editorial System(tm) for Geomorphology

Manuscript Draft

Manuscript Number: GEOMOR-445R2

Title: Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment

Article Type: Research Paper

Section/Category:

Keywords:

Corresponding Author: Mr. Yannick Thiery, PhD

Corresponding Author's Institution: UMR 6554 CNRS LETG-Geophen

First Author: Yannick Thiery, PhD

Order of Authors: Yannick Thiery, PhD; Jean-Philippe Malet, Researcher; Simone Sterlacchini, Researcher; Anne Puissant, Lecturer; Olivier Maquaire, Professor

Manuscript Region of Origin:

Abstract: Statistical assessment of landslide susceptibility has become a major topic of research in the last decade. Most progress has been accomplished on producing susceptibility maps at meso-scales (1:50,000-1:25,000). At 1:10,000 scale, which is the scale of production of most regulatory landslide hazard and risk maps in Europe, few tests on the performance of these methods have been performed. This paper presents a procedure to identify the best variables for landslide susceptibility assessment through a bivariate technique (weights of evidence, WOE) and discusses the best way to minimize conditional independence (CI) between the predictive variables. Indeed, violating CI can severely bias the simulated maps by over- or under-estimating landslide probabilities. The proposed strategy includes four steps: (i) identification of the best response variable (RV) to represent landslide events, (ii) identification of the best combination of

predictive variables (PVs) and neo-predictive variables (nPVs) to increase the performance of the statistical model, (iii) evaluation of the performance of the simulations by appropriate tests, and (iv) evaluation of the statistical model by expert judgment. The study site is the north-facing hillslope of the Barcelonnette Basin (France), affected by several types of landslides and characterized by a complex morphology. Results indicate that bivariate methods are powerful to assess landslide susceptibility at 1:10,000 scale. However, the method is limited from a geomorphological viewpoint when RVs and PVs are complex or poorly informative. It is demonstrated that expert knowledge has still to be introduced in statistical models to produce reliable landslide susceptibility maps.

Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment.

Y. Thiery [a,*], J.-P. Malet [a], S. Sterlacchini [b], A. Puissant [c], O. Maquaire [a]

[a] UMR 6554 CNRS, LETG-Geophen, University of Caen-Basse Normandie, Esplanade de la Paix, BP 5183, F-14032 Caen Cedex, France

[b] National Research Council, Institute for the Dynamic of Environmental Processes,,Piazza della Scienza, 1, I-20126 Milano, Italy

[c] UMR 2795 CNRS, IDEES-GeoSyscom, University of Caen-Basse Normandie, Esplanade de la Paix, F-14032 Caen Cedex, France

Abstract

Statistical assessment of landslide susceptibility has become a major topic of research in the last decade. Most progress has been accomplished on producing susceptibility maps at meso-scales (1:50,000-1:25,000). At 1:10,000 scale, which is the scale of production of most regulatory landslide hazard and risk maps in Europe, few tests on the performance of these methods have been performed. This paper presents a procedure to identify the best variables for landslide susceptibility assessment through a bivariate technique (weights of evidence, WOE) and discusses the best way to minimize conditional independence (CI) between the predictive variables. Indeed, violating CI can severely bias the simulated maps by over- or under-estimating landslide probabilities. The proposed strategy includes four steps: (i) identification of the best response variable (RV) to represent landslide events, (ii) identification of the best combination of predictive variables (PVs) and neo-predictive variables (nPVs) to increase the performance of the statistical model, (iii) evaluation of the performance of the

simulations by appropriate tests, and (iv) evaluation of the statistical model by expert judgment. The study site is the north-facing hillslope of the Barcelonnette Basin (France), affected by several types of landslides and characterized by a complex morphology. Results indicate that bivariate methods are powerful to assess landslide susceptibility at 1:10,000 scale. However, the method is limited from a geomorphological viewpoint when RVs and PVs are complex or poorly informative. It is demonstrated that expert knowledge has still to be introduced in statistical models to produce reliable landslide susceptibility maps.

Keywords: Landslide, Susceptibility assessment, GIS, Statistical modeling, Weights of evidence, Expert knowledge, French Alps

---

* Corresponding author. Tel.: + 33(0)3 90 24 09 28
E-mail address: thiery@equinoxe.u-strasbg.fr

1. Introduction

Assessing landslide hazard and risk with a minimum set of data, a reproducible methodology and GIS techniques, is a challenge for earth-scientists, government authorities and resource managers (Glade and Crozier, 2005). Landslide hazard assessment (LHA) estimates the probability of occurrence of landslides in a territory within a reference period (Varnes, 1984; Fell, 1994; van Westen et al., 2006). It is deduced from information on (i) landslide susceptibility expressed as the spatial correlation between predisposing terrain factors (slope, land use, superficial deposits, etc.) and the distribution of observed landslides in a territory (Brabb, 1984; Crozier and Glade, 2005) and, (ii) the temporal dimension of landslides related to the occurrence of triggering events (rainfalls, earthquakes, etc.). In most cases, landslide frequencies are difficult to obtain due to the absence of historical landslide records. Therefore, LHA is most of the time restricted to landslide susceptibility assessment (LSA) which is considered as a 'relative hazard assessment', and does not refer to the time dimension of landslides (Sorriso Valvo, 2002). Landslide susceptibility maps can be obtained by two

54 categories of methods: (i) direct approaches based on expert knowledge of the target area, and
55 (ii) indirect approaches based on statistical algorithms.

56 The direct approaches are based on expert knowledge about the relation between the
57 occurrences of landslides and their hypothesized predisposing factors. The approach
58 necessitates the definition of expert rules leading to different susceptibility degrees (Soeters
59 and van Westen, 1996). In France, the official methodology to assess landslide susceptibility
60 and hazard is based on direct approaches. The methodology, called 'Plans de Prévention des
61 Risques' (MATE/MATL, 1999) has been applied at 1:10,000 scale.

62 The main concept of the indirect approaches is that the controlling factors of future landslides
63 are the same as those observed in the past (Carrara et al., 1995). Indirect approaches are based
64 on statistical conditional analyses and on the comparisons of landslide inventories and
65 predisposing terrain factors. The methods are applied at the scale of the terrain unit (TU)
66 corresponding to a portion of hillslope possessing a set of predisposing factors, which differs
67 from that of the adjacent units with definable boundaries (Hansen, 1984; Carrara et al., 1995).
68 Indirect approaches predict landslide distribution (the response variable, RV) through a set of
69 *a priori* independent terrain factors (the predictive variables, PVs).

70 Several bivariate (certainty factors and weights of evidence) or multivariate (logistic
71 regression and discriminant analysis) approaches were developed for landslide susceptibility
72 mapping. A synthesis of the available methods, their applicability and drawbacks, can be
73 found in Yin and Yan (1988), Carrara et al. (1995), Chung et al. (1995), Soeters and
74 van Westen (1996), Atkinson and Massari (1998), Aleotti and Chowdury (1999), Guzetti et
75 al. (1999), Clerici et al. (2002), Dai et al. (2002), van Westen (2004) and van Westen et al.
76 (2006). In the scientific community it is commonly admitted that statistical analyses are more
77 appropriate for susceptibility zoning at meso-scales (1:50,000 to 1:25,000) because of their

78 potential to minimize expert subjectivity (Soeters and van Westen, 1996; van Westen et al.,
79 2006).

80 Although the bivariate approaches are considered as more robust and flexible (van Westen et
81 al., 2003; Süzen and Doruyan, 2004), they present some limitations:

82 (i) The tendency to over-simplify the (input) thematic data (e.g. predisposing factors) that
83 condition landslides, by taking only what can be relatively easily mapped or derived
84 from a DTM (van Westen et al., 2003, 2006).

85 (ii) The large sensitivity to the quality and accuracy of the thematic data, e.g., imprecision
86 and incompleteness of landslide information, and limited spatial accuracy of information
87 on the predisposing factors (Guzzetti et al., 2006). Application of the methods is
88 relatively limited at large scales because most of thematic data are available only at
89 meso-scales (1:50,000 to 1:25,000). Especially for most mountain areas a discrepancy
90 remains between the scale of available data and the scale of landslide occurrence. For
91 instance, geological maps and land-use maps are available only at scales from 1:50,000
92 to 1:25,000 for most parts of the French Territory; also, only digital terrain models with
93 a planimetric resolution of 50 m and a vertical accuracy of 2 to 3 m are available. These
94 input data are not adapted to the analysis of landslide susceptibility at 1:10,000 scale
95 (Thiery et al., 2003, 2004).

96 (iii) The singularity of predisposing factors for each landslide type, which forces us to
97 analyse them individually in order to have distinct susceptibility maps (Atkinson and
98 Massari, 1998; Kojima et al., 2000; van Westen et al., 2006).

99 (iv) The number of landslide events to incorporate in the statistical model in relation to the
100 size of the study area (Bonham-Carter, 1994; Begueria and Lorente, 1999; van den
101 Eeckaut et al., 2006).

(v) The use of statistically independent predictive variables in the application of bivariate methods. When the influence of a combination of predictive variables on the response variable is evident, the weight associated to each thematic factor is calculated independently and combined in a unique equation (Agterberg et al., 1993; Bonham-Carter, 1994). The probabilities computed with this equation may be different from those calculated directly from the input data. Therefore, applying the method requires to assume conditional independence (CI) of the dataset (Bonham-Carter et al., 1989; Agterberg et al., 1993; van Westen, 1993; Agterberg and Cheng, 2002; Thiart et al., 2003).

(vi) The absence of expert opinions if the method is applied by GIS experts and not by earth-scientist. In other words, the model should give satisfactory results in term of degree of fit, but should also correspond to the 'real world' (van Westen et al., 2003, 2006).

Some procedures were proposed to overcome these limitations and increase the robustness of landslide susceptibility assessments with indirect approaches through: (i) proper validation and reduction of simulation uncertainty (Chung and Fabbri, 2003; Chung, 2006; Guzzetti et al., 2006; van den Eeckaut et al., 2006), (ii) reduction of the costs of data acquisition (Greco et al., 2007), and (iii) introduction of expert knowledge to the statistical models used (van Westen et al., 2003).

Hence, the aim of this work is to ascertain a reproducible procedure to estimate landslide susceptibility with a bivariate approach at 1:10,000 scale in a complex mountainous environment, while limiting the collection of landslide and thematic data. The procedure adopted for this research includes four steps:

(i) Identification of the best way to calculate landslide probabilities based on the characteristics of the landslide inventory.

(ii) Identification of the most relevant combination of predisposing terrain factors avoiding conditional dependence.

(iii) Evaluation of the degree of model fit by statistical tests and comparisons with the landslide inventory.

(iv) Evaluation of the best indirect susceptibility map in comparison with a direct susceptibility map.

The procedure was applied to the north-facing hillslope of the Barcelonnette Basin (South French Alps) affected by several landslide types (Maquaire et al., 2003; Thiery et al., 2005; Malet et al., 2005).

2. Geomorphological settings

2.1. Geomorphology of the Barcelonnette Basin

The Barcelonnette Basin is representative of climatic, lithological, geomorphological and land-use conditions observed in the South French Alps, and is highly affected by landslide hazards (Flageollet et al., 1999). It is situated in the dry intra-Alpine zone, characterized by a mountain climate with a Mediterranean influence. Highly variable rainfall amounts (400 to 1300 mm yr^{-1}) occur with intense storms during summer and autumn. However, as pointed out by Flageollet et al. (1999), landslides there are not controlled only by climatic conditions; slope instability can occur after relatively dry periods whether or not preceded by heavy rainfalls.

The test site extends over an area of about 100 km^2. Located on the north-facing hillslope (Fig. 1), it is characterized by a large variety of active landslides and is representative of the environmental conditions observed in the Barcelonnette Basin. The Ubaye River depicts the northern boundary, while the Sauze torrent delimits the western boundary; the southern and

150 eastern boundaries are represented by high crests of limestones and sandstones. The test site
151 can be subdivided into two geomorphological units separated by a major fault in a north/south
152 direction. The eastern unit is dominated by allochthonous sandstones outcrops, while the
153 western unit is composed of autochthonous Callovo-Oxfordian marls (BRGM, 1974;
154 Flageollet et al., 1999; Maquaire et al., 2003).

155 The eastern unit (ca. 40 km^2) is drained by the Abriès torrent which cuts an asymmetric valley
156 in highly fractured sandstones. The gentle slopes there (10-30°) are covered by moraine
157 deposits of 2 to 15 m thick and by coniferous forests or grasslands (Fig. 2); these slopes are
158 affected by shallow rotational or translational slides triggered by the undercutting of torrents.
159 In contrast, the steep slopes (30-70°) are characterized by bare soils and affected by rockfalls
160 on sandstones.

161 The western unit (ca. 60 km^2), drained by four main torrents, presents an irregular topography
162 of alternating steep convex slopes, planar slopes and hummocky slopes. The steepest convex
163 slopes (>35°) are carved in black marl outcrops, and are very commonly gullied into badlands,
164 or affected by rock-block or complex slides (Malet et al., 2005). The planar slopes (5-30°)
165 composed of thick moraine deposits (from 6 to 20 m), are very often cultivated and affected
166 by rotational or translational slides. The hummocky slopes are generally covered by forests
167 and/or natural grasslands (Fig. 2), and affected by large relict landslides and/or surficial soil
168 creep. Most landslides within the western unit are located along streams or on gentle slopes,
169 where the contact of moraine deposits and black marls creates a hydrological discontinuity
170 favourable for slope movements.

171

172 *2.2. Landslide data*

173 A landslide inventory was compiled at 1:10,000 scale through air photo-interpretation, field
174 surveys and analysis of literature in years 2002 and 2003 by a geomorphologist (Thiery et al.,
175 2003, 2004). Air-photo interpretation was carried out on 1:25,000-scale photographs (year
176 2000) issued from the French Geographical Institute. Fieldwork was carried out between July
177 2002 and July 2003 to complete the photo-interpretation. To reduce uncertainty linked to an
178 expert in charge of mapping (Ardizzone et al., 2002; Wills and Mc Crinck, 2002), two
179 degrees of confidence were defined for the photo-interpretation and information of available
180 literature (landslide recognition or not), while three degrees of confidence (high, medium and
181 low) were distinguished for the field survey. A mapping confidence index (MCI) in three
182 classes (high, medium and low) was derived. Three hundred fourteen landslides were
183 recognized, with 66% classified with a high MCI , 27% with a medium MCI and 7% with a
184 low MCI. Among the 207 landslides with a high MCI, 10% are considered as relict, 8% are
185 considered as latent, and 82% are considered as active. The active landslides can be grouped
186 in three types (Table 1) according to the typology of Dikau et al. (1996).

187 Figs. 3 and 4 present the morphology and morphometric/environmental characteristics of the
188 landslides. Shallow translational slides are relatively small and mainly located on steep slopes
189 along streams. They occur on the weathered bedrock or in moraine deposits. Rotational slides
190 are located along streams but more on gentle slopes than the shallow translational slides. They
191 occur principally in moraine deposits or at the contact with the bedrock. Translational slides
192 are located more on gentle slopes at the contact with the bedrock, and their sizes are very
193 variable (Table 1).

194 The boundaries of active landslides were classified into two zones and digitized: (i) the
195 landslide triggering zone (LTZ) and (ii) the landslide accumulation zone (LAZ, Fig. 3). The
196 geometrical (perimeter, area, and maximal length) and geomorphological characteristics
197 (typology and state of activity) were stored in a GIS database.

198 As the aim of this study is to locate areas prone to failures, only the LTZ of active landslides
199 were introduced in the analysis (Atkinson and Massari, 1998; van den Eeckhaut et al., 2006).
200 In statistical models, the total area of landslides (van Westen et al., 2003) or only the
201 triggering area can be used to compute probabilities of landsliding (Chung and Fabbri, 2003;
202 Remondo et al., 2003). According to the characteristics of the landslides, especially their run-
203 out distances, a severe bias can occur when the landslide accumulation zone is taken into
204 account in the model. Indeed, several classes of input data may be included in the probability
205 calculation process, while in reality they were not the most important controlling factors.
206 Therefore, Atkinson and Massari (1998), Sterlacchini et al. (2004), and van den Eeckhaut et
207 al. (2006) proposed to use only one cell at the centre of the triggering zone. This procedure
208 offers some advantages because it does not take into account the landslide boundaries and it
209 does not attribute a too large influence to the largest landslides which exhibit more diversity
210 in predisposing factors. However, if the results based on one cell at the centre of the triggering
211 zone can be satisfactory, the final probabilities are not necessarily representative of the
212 predisposing conditions at the onset of the landslide. Defining the most appropriate part of the
213 landslide to compute the probabilities is therefore a prerequisite to understand how it
214 influences the model results.

215

216 *2.3. Landslide predisposing factors*

217 The statistical analysis of the landslide inventory has outlined the main predisposing factors
218 (predictive variable) to introduce in the statistical model. The thematic data (Table 2) are
219 derived from (i) available national databases, (ii) air-photo interpretation analyses, (iii)
220 satellite imagery analyses, and (iv) field surveys. The DTM (10-m resolution) was constructed
221 by the kriging interpolation applied to a network of triplets, obtained from the digitisation of

222 contour lines in 1:25,000-scale topographic maps which were enlarged by the French
223 Geographical Institute into 1:10,000 scale. Its accuracy is of about ±1 m for the horizontal
224 component, and ±2 to 10 m for the vertical component, depending on relief.

225 The slope gradient map and the slope curvature map were derived from the DTM. The
226 lithological map is based on the main lithological units described in a geological map
227 produced by the French Geological Survey (BRGM, 1974) at 1:50,000 scale, and was
228 completed by fieldwork. The surficial formation map was obtained by the segmentation of the
229 landscape into homogeneous macro-areas closely associated with sediment facies (van
230 Westen, 1993). The surficial formation thickness map was derived from direct observations of
231 outcrops along streams and steep slopes. The land-use map was produced by the analysis of a
232 Landsat ETM+ image (year 2000) fused with a SPOT-P image (year 1994); the boundaries of
233 homogeneous land-use units were corrected by air-photo interpretation.

234

235 *2.4. Direct landslide susceptibility map*

236 The direct landslide susceptibility map was elaborated with the French legal procedure for
237 landslide hazard and risk at 1:10,000 scale (MATE/MATL, 1999; Leroi, 2005). This
238 methodology requires a global overview of the area to identify sectors with homogeneous
239 environmental characteristics for each landslide type. The methodology advises us to take into
240 account the possibilities of landslide development for the forthcoming one hundred years.
241 Four degrees of susceptibility were defined. The expert rules used to define the direct
242 susceptibility classes are detailed in Table 3.

243

244 **3. Methodology and strategy**

3.1. Weights-of-evidence (WOE): background

3.1.1. WOE method

Weights-of-evidence (WOE) is a quantitative 'data-driven' method used to combine datasets. The method, first applied in medicine (Spiegelhater and Kill-Jones, 1984) and geology (Bonham-Carter, 1994), uses the log-linear form of the Bayesian probability model to estimate the relative importance of evidence by statistical means. This method was first applied to the identification of mineral potential (Bonham-Carter et al., 1990) and then to landslide susceptibility mapping (van Westen, 1993; van Westen et al., 2003; Süzen and Doruyan, 2004).

Prior probabilities (PriorP) and posterior probabilities (PostP) are the most important concepts in the Bayesian approach. PriorP is the probability that a TU (terrain unit) contains the RV (response variable) before taking PVs (predictive variables) into account, and its estimation is based on the RV density for the study area. This initial estimate can be modified by the introduction of other evidences. PostP is then estimated according to the RV density for each class of the PV. The model is based on the calculation of positive W^+ and negative W^- weights, whose magnitude depends on the observed association between the RV and the PV.

$$W^{+} = \ln \frac{P(B \mid \mathrm{RV})}{P(B \mid \overline{\mathrm{RV}})} \quad (1)$$

$$W^{-} = \ln \frac{P(\overline{B} \mid \mathrm{RV})}{P(\overline{B} \mid \overline{\mathrm{RV}})} \quad (2)$$

263 In Eqs. (1) and (2), B is a class of the PV and the overbar sign '¯' represents the absence of
264 the class and/or RV. The ratio of the probability of RV presence to that of RV absence is
265 called odds (Bonham-Carter, 1994). The WOE for all PVs is combined using the natural
266 logarithm of the odds (logit), in order to estimate the conditional probability of landslide
267 occurrence. When several PVs are combined, areas with high or low weights correspond to
268 high or low probabilities of presence of the RV.

269

270 *3.1.2. Hypothesis of the WOE method*

271 As mentioned by Bonham-Carter (1994), the results of the WOE method are strongly
272 dependent on the number of events introduced in the model (e.g. on the estimation of
273 probabilities) and on the quality of the landslide inventory map. Therefore, probabilities are
274 very low if the area is characterized by rare events, and the results have to be interpreted
275 cautiously. Nevertheless, if the study area is covered by reasonable samples of events, the
276 estimated weights can be stable and realistic.

277 The WOE method requires the assumption that input maps are conditionally independent. To
278 meet this need, many statistical tests may be used (e.g., χ^2-test, omnibus test, and new
279 omnibus test). A detailed review of the performance of these tests can be found in Agterberg
280 and Cheng (2002) and Thiart et al. (2003). In case of violation of conditional independence,
281 PVs which are dependent can be combined into a neo-variable (nPV) which is then used in
282 the WOE method (Thiart et al., 2003). The weighted-logistic-regression method (WLR) may
283 also be used to bypass the violation of conditional independence. However, if the density of
284 the RV is low, this method severely underestimates PostP, and a number of the RV smaller
285 than the observed value can be predicted (Thiart et al., 2003). Consequently, specific

286 procedures have to be used on large areas characterized by a low density of the RV (Begueria
287 and Lorente, 1999; van den Eeckhaut et al., 2006).

288

289 *3.2. Employed methodology*

290 The employed methodology uses the main steps described by van Westen et al. (2003) and
291 Guzzetti et al. (2006), i.e.: (i) aptitude of thematic data to construct a model, (ii) evaluation of
292 the uncertainty level of probabilities, (iii) determination of the degree of model fit
293 (performance) to an indirect landslide susceptibility map, and (iv) evaluation of the indirect
294 landslide susceptibility map in comparison with a direct susceptibility map.

295 The first three steps were tested on a 'sampling area' of the study site (north-facing hillslope
296 of the Barcelonnette Basin) characterized by the occurrence of the three types of landslides
297 (Fig. 1). This test area extends over about 11 km^2 and is representative of the western and
298 eastern terrain units described previously. The upper parts of the hillslopes were not included
299 in the 'sampling area' because the environmental conditions are not representative of the
300 landslides introduced in the analysis.

301 The probabilities of future landslide occurrence are calculated for each landslide type (only
302 LTZs are introduced in the analysis) and a susceptibility map is created after the classification
303 of PostP. Susceptibility classes were compared to the observed LTZs in the 'sampling area'. If
304 results were satisfactory, the statistical model was applied to the whole area with the same
305 procedure (Fig. 5). Then, the final indirect landslide susceptibility map was assessed with the
306 direct landslide susceptibility map with a confusing matrix and several statistical accuracy
307 tests. Thus, a careful confrontation with a reference map was performed at each step. The
308 statistical model was implemented in ArcView 3.2® through the ArcSDM extension (Kemp
309 et al., 2001), and the size of the calculation cell was 10 m.

310

311 *3.2.1. Identification of the response variable (RV)*

312 Bayesian models are very sensitive to the number and quality of the RV. Over large areas
313 characterized by complex thematic data, it can be very difficult to identify LTZs with high
314 confidence. To deal with these limitations, the first two steps of the procedure are: (i) to
315 identify the minimum number of cells representing the variability of the predisposing factors
316 within LTZs, and (ii) to identify the best spatial location of cells to represent the variability of
317 the predisposing factors within LTZs. For each landslide type, the same number of cells was
318 introduced at each calibration phase. The initial number of cells in the LTZs examined in this
319 study is 460.

320 The minimal number of cells to introduce in the model was estimated by a random sampling
321 (10 to 100%) of the LTZ cells of each landslide type. The best spatial location of cells was
322 estimated by selecting several cells' locations within the LTZs (Table 4). The computations
323 were performed with a set of four *a priori* 'constant' thematic maps of PVs (slope gradient,
324 surficial deposits, lithology, and land use). A landslide susceptibility map was then produced
325 for each combination. The PostP distribution was analysed by expert judgment to define
326 susceptibility classes. In former studies, the number of classes varied from two (e.g. stable
327 and unstable; Begueria and Lorente, 1999) to six (null, very low, low, moderate, high, and
328 very high susceptibility; Chacón et al., 2006). In this study, landslide susceptibility was
329 classified into four (null, low, moderate, and high) for comparison to the direct landslide
330 susceptibility map with the four classes. The relative error ξ was computed to evaluate the
331 performance of the simulations:

332 $$\xi = \frac{O_L - P_L}{O_L} \quad (3)$$

where O_L is the number of the observed landslide cells representing the LTZ of active landslides, and P_L is the number of the predicted landslide cells with the high susceptibility class. If the relative error decreases with the introduction of a RV, this RV is retained for the next simulation step (Fig. 5).

3.2.2. Identification of the predictive variables (PVs)

The performance of the PVs introduced successively in the statistical model was evaluated in terms of CI violation and distribution of PostP for each landslide type. Computations were performed with the best RV dataset identified previously. The procedure is as follows:

(i) Selection of the best PV dataset by expert judgement which takes into account the predisposing factors and classes associated with each landslide type;

(ii) Analysis of CI violation between each PV and the RV. As the χ^2-test is very sensitive to the density of the RV introduced in the model (Thiart et al., 2003) and may increase the measure of the dependence between two PVs by 25 to 30% (Pistocchi et al., 2002; Dumolard et al., 2003), the Cramer's V coefficient (Kendall and Stuart, 1979) is calculated. The Cramer's V is considered as the more robust association test because of its possibility to assess large and complex contingency tables (Howell, 1997). The coefficient provides a standardized measure in the range [0-1]; the closer $V \rightarrow 1$, the stronger is the association between two PVs.

(iii) Exploration of the structure of the association between PV classes and the RV by a multiple correspondence analysis (MCA), and definition of the most significant classes of a PV to represent landslide occurrences.

(iv) Introduction of a neo-variable (nPV) with geomorphological meaning (van Westen et al., 2003) in the statistical model by combining PVs causing CI violations.

(v) Finally, the performance of each PV and nPV is assessed by introducing the variables iteratively in the statistical model. If the relative error does not decrease despite the addition of a PV or an nPV, the simulation is rejected; whereas, if the relative error decreases, the simulation is accepted.

3.2.3. Evaluation of performance of the indirect susceptibility maps

The performance of the indirect susceptibility maps was assessed for the total study area with the best combination of PVs and nPVs (Figs. 1 and 5). Both statistical and expert evaluations were performed successively.

First, the weights obtained for the classes of the best PVs and nPVs are applied to the total study area (Figs. 1 and 5) and the susceptibility classes were defined with the same thresholds in the cumulative curves. The degree of model fit was evaluated by analysing the ξ value for all the LTZs observed in the total study area. If ξ is low (<0.3), the statistical model is considered as robust. Then, the confidence of PostP was evaluated by the Student-t test. This test uses the variance of PostP to create a normalized value to estimate the certainty of the calculation with the null hypothesis H_0: PostP = 0. The normalized value has to be equal or larger than 1.64 to have a certainty calculation of 95% (Bonham-Carter, 1994; Davis, 2002).

Second, the indirect susceptibility map was compared with the direct susceptibility map. Because the direct susceptibility map had been produced by the French Official Method of Landslide Risk Zoning (MATE/METL, 1999) independently of the landslide types, a unified indirect susceptibility map was produced by combining the indirect susceptibility maps obtained for the three landslide types. The four classes of the indirect susceptibility maps were merged, and for each cell, more weight was systematically given to the higher susceptibility class (Fig. 8). Confusion matrices were calculated and several statistical tests

were performed for the direct and unified indirect susceptibility maps (Tables 5 and 6). The Kappa (*K*) coefficient was used to assess the improvement of the model predictions over chance (Table 6). A *K* value of 1 is equivalent to a perfect agreement between the model and the reference map. *K* values higher than 0.4 signify a good statistical agreement between maps (Fielding and Bell, 1997).

4. Results

4.1. Best response variable

The minimum number of cells representing the variability of the predisposing factors within the LTZs was identified from the 460 cells. The relation between the number of LTZs cells introduced in the model and ξ for each landslide type is presented in Fig. 7. A threshold comparable to 50% of the 460 cells was identified to stabilize ξ for the 'sampling area', and the simulations with RV-3 to RV-7 were performed with the 230 cells. Table 4 indicates that the simulations with RV-2 and RV-3 are not acceptable, confirming that using only one or a few cells around the centre of a LTZ mass underestimates PriorP and PostP. Table 4 also indicates the influence of LTZ sizes on the results, and highlights that the best results are obtained with the use of the cells representing the most frequent combination of PVs observed in LTZs (RV-7).

4.2. Best predictive variables

Statistical tests indicate CI violation between the PVs. As an example, the values of the χ^2-test and the Cramer's *V* coefficient for the translational slides are detailed in Table 7. The

403 Cramer's *V* coefficient indicates a low association between the variables except for SLO-CUR
404 and SLO-SF. The correlation SLO-CUR is mainly related to the location of RV-7 cells on
405 slopes between 15° and 35°, which cover more or less 50% of the 'sampling area' and present
406 planar slopes. Therefore, the information contained in these two PVs is redundant and
407 combining these variables has no geomorphological meaning. Consequently, the PV CUR
408 was not introduced in the statistical model. In contrast, the combination of variables with a
409 geomorphological meaning (for instance SLO and SF) was introduced.

410 The first four axes of the MCA (multiple correspondence analysis) explain 40.5%, 49.3% and
411 46.0% of the total variance for the shallow translational slides, rotational slides and
412 translational slides, respectively. Despite the low contribution of each axis (<20%) on the
413 cumulated variance, some useful information is still highlighted by the MCA. For example,
414 the axes F1, F2 and F3 of the translational slides confirm the relation between SLO and the
415 surficial formations (SF and TSF). Thus, the MCA gives some indications on the possible
416 combination of classes for each PV, and allows us to justify the definition of an nPV with
417 both a geomorphological meaning and a low redundancy of information. Table 8 summarizes
418 the results of the MCA for the three landslide types. Fig. 8 details the cumulative curves
419 associated with each WOE simulations and the different thresholds to define the four
420 susceptibility classes for each landslide type. Fig. 9 presents the susceptibility maps obtained
421 for the shallow translational slides. Simple geomorphological information given by the nPV
422 increases the performance of the models. For example, for the shallow translational slides, the
423 best simulation carried out with the non-combined PVs (SLO, FS, LIT, and LAD) is
424 characterized by a ξ value of 0.45 (Table 6), while the best simulation with the introduction of
425 nPV-1 (which combines slope gradient classes and surficial formation types, Table 9) is
426 characterized by a ξ value of 0.14 (Table 9). For the simulations performed in the 'sampling

427 area', ξ values are 0.18, 0.16, and 0.14 for the shallow translational slides, rotational slides,
428 and translational slides, respectively (Table 9).

429

430 *4.3. Evaluation of indirect susceptibility maps*

431 Fig. 10 presents the indirect susceptibility maps for each landslide type obtained by applying
432 the PostP of the 'sampling area' to the total study area. The maps show a good agreement
433 with the landslide inventory map and are characterized by ξ values of 0.22, 0.25 and 0.23 for
434 the shallow translational slides, the rotational slides, and the translational slides, respectively
435 (Table 10). The surfaces of high, moderate and low susceptibility are 4.9 km^2, 1.6 km^2 and 1.6
436 km^2 for the shallow translational slides, 12.3 km^2, 5.1 km^2 and 6.3 km^2 for the translational
437 slides, 3.8 km^2, 2.2 km^2 and 3.2 km^2 for the rotational slides, and 12.3 km^2, 5.1 km^2 and 6.3
438 km^2 for the translational slides, respectively. The certainty test indicates a percentage of
439 presence of the high susceptibility class in the confidence zone of 70.8%, 88.7% and 87.5%
440 for the shallow translational slides, rotational slides, and translational slides, respectively.
441 Consequently the high susceptibility classes simulated with the statistical models
442 incorporating an nPV are relevant from a statistical viewpoint.

443 The unified indirect susceptibility map (Fig. 11) was then compared to the direct
444 susceptibility map (Fig. 12). The former map identifies 17.7 km^2, 5.8 km^2 and 6.9 km^2 of the
445 high, medium and low susceptibility classes, respectively (Fig. 11). The confusion matrix
446 (Table 11) indicates a good accuracy between the direct and indirect maps, especially for the
447 high susceptibility class. Fig. 13 presents the observed differences between the two maps
448 concerning the high susceptibly class.

449

5. Discussion

The proposed methodology to assess landslide susceptibility at 1:10 000 scale is based on a bivariate method calibrated on a 'sampling area' and validated on a larger area. To obtain a robust and reproducible procedure, simple and easy-to-obtain thematic data with a high cost-benefit ratio were used. The thematic maps introduced in the statistical model represent slope gradient, slope curvature, surficial formations, thickness of surficial formations, lithology, land use and streams. Our work indicates that introducing only simple PVs in the statistical model does not satisfactorily recognise landslide-prone areas in a complex environment. Therefore, the concept of nPV, the use of the main set of predisposing factors for one landslide type, was employed. In our case this set is essentially represented by the combination of the thematic classes of slope gradients and surficial deposits. An nPV is identified by analysing the structure of the relationships between the landslide types, slope gradients and surficial formations. The nPV significantly increases the performance of the three statistical models, as pointed out by the decrease of the ξ value from 0.45 to 0.14 for the shallow translational slides, 0.43 to 0.16 for the rotational slides, and 0.40 to 0.18 for the translational slides. Evaluation of the statistical model for the total study area shows good agreement among the indirect susceptibility map, the landslide inventory map, and the direct susceptibility map. However, to obtain a good agreement, several considerations have to be pointed out:

(i) Our indirect susceptibility maps represent better the high susceptibility class than the low to moderate susceptibility classes. Tables 10 and 11 confirm the good agreement of the indirect susceptibility map with the landslide inventory map and the direct susceptibility map for the high susceptibility class. The indirect susceptibility maps underestimate the surfaces of the low and moderate susceptibility classes with K values of 0.03 and 0.08, respectively. These disagreements are explained by the methodology

used to produce the direct and indirect susceptibility maps. On the one hand, rules relying on expert judgments can take into account (i) some subtle changes in specific areas which modify the degree of susceptibility, and (ii) the possibility of spatial evolution of landslides. On the other hand, statistical models were developed in our study to recognize areas favourable for active LTZs. The calculation processes of such models are based on binary evidences and are optimized to recognize areas with identical environmental characteristics, and the procedure of calibration/validation of the models is dependent on the thresholds observed on the simulated cumulative curves (Begueria, 2006; van den Eeckhaut et al., 2006). If this classification/validation procedure is employed, some potentially landslide-prone areas may be overestimated or underestimated (Begueria, 2006), and consequently the low and moderate susceptibility classes are not very well identified on the cumulative curve.

(ii) Our indirect susceptibility maps may not take some portion of terrain into account. For instance, in our study, the portions of terrain with slope gradients lower than 15° are always considered with a low or null susceptibility, although some of such areas are prone to landsliding. This discrepancy may be explained by the analysis used to select the best RV (RV-7) which mathematically increases the weights of the PV combination corresponding to the LTZs, and by the underestimation of PostP for these slope gradients because only a few LTZs are located on these slopes.

(iii) On a more general viewpoint, the 'sampling area' has to be selected carefully. Indeed, if the 'sampling area' is not sufficiently representative of the environmental conditions of the total study area, calculations of PriorP and PostP are biased. If the study area is sufficiently large, a sensitivity analysis on several 'sampling areas' with different sizes and shapes is recommended in order to select the more appropriate area which represents the total study area (Greco et al., 2007). In our case, the study area has a

complex topography with two distinct parts and several landslide types. Therefore, the selection of the 'sampling area' was based on geomorphological knowledge of the site.

(iv) Statistical models are very sensitive to the type and number of landslide cells. A conceptual model has therefore been created for each landslide type, because each type is controlled by a specific combination of predisposing factors. Furthermore, the quality of the indirect susceptibility maps depends on the selection of relevant cells representing the variability of the environmental factors (Greco et al., 2007).

(v) Statistical models are also very sensitive to the thematic data of environmental factors, and to their potential conditional dependence. Regarding CI violation, the results of the χ^2-test and the value of the V coefficient have to be interpreted with caution, because a few cells can severely bias the results (Dumolard et al., 2003). These tests are just informative and they cannot be used in rigorous terms (Pistocchi et al., 2002). Therefore, instead of not incorporating the cells posing some problems or decreasing the total number of RV cells, the proposed procedure intends to combine some classes of the PVs which are conditionally dependent. Indeed, decreasing the number of RV cells could modify the stability of the model as demonstrated previously. A robust procedure to follow is to combine an expert judgment with the χ^2-test and the V coefficient in a multiple correspondence analysis, in order to identify the classes of PVs violating CI and select the classes of PVs to be combined with an nPV with a geomorphological meaning. As mentioned by van Westen et al. (2003, 2006), expert judgment is very important in the conception of the statistical model to guide thematic maps towards geomorphological landslide evidences. Regarding the minimum set of thematic maps, the different statistical tests used in our study stress the difficulty to map landslide susceptibility at 1:10,000 scale using only a few variables. Other data sources such as a more detailed soil thickness map or detailed structural maps (fault map and tectonic

map) should be used in order to obtain more accurate results. Nevertheless, at this scale of work and for a large and complex environment, these variables are extremely difficult to measure because of their high spatial variability. Therefore, they have been often neglected in susceptibility assessment.

The proposed procedure follows the guidelines suggested by van Westen et al. (2003) and Guzzetti et al. (2006) for the validation of indirect susceptibility maps. Guzetti et al. (2006) proposed a set of criteria for ranking and comparing the quality of landslide susceptibility assessments, i.e., the quality of the input data and the use of different statistical tests. In terms of these criteria, the susceptibility maps obtained with the procedure used in this study have the highest quality (level 7).

6. Conclusion

This study has demonstrated the necessity of using specific and adapted procedures for indirect landslide susceptibility assessment by bivariate methods, especially at 1:10,000 scale, for complex environments with some uncertainty in collected landslide characteristics. The proposed procedure, based on a reduced number of thematic data and a 'sampling area', consists of four steps. First, the best response variable RV (e.g. landslide inventory) to be introduced in the statistical model is defined. This variable may vary according to the landslide type and the environmental characteristics of the study area. Second, the best PVs (e.g. terrain predisposing factors) to be used in the statistical model are identified by minimizing conditional dependence on the basis of statistical tests. The structure of the statistical relation between RV and PV is studied through multiple correspondence analyses to identify the class of PVs influencing the location of landslides. Based on the results, neo-predictive variables (nPVs) with geomorphological meanings are proposed, and introduced in the statistical models. Third, the performance and confidence associated with the simulations

549 are evaluated by statistical tests and expert knowledge. Fourth, more appropriate thematic
550 data and weights identified on the 'sampling area' are applied to the total study area. The
551 results are compared to a direct landslide susceptibility map through a confusion matrix.

552 The procedure was applied successfully to the north-facing hillslope of the Barcelonnette
553 Basin. The indirect and direct susceptibility maps are quite similar for the high susceptibility
554 class with a high classification rate and a good Kappa (*K*) coefficient.

555 This study has demonstrated that the use of a 'sampling area' correctly representing the
556 geomorphology of a larger area, combined with the use of neo-predictive variables, is
557 sufficient to calibrate a bivariate statistical model for landslide susceptibility assessment. This
558 study reinforces the use of bivariate statistical models based on both expert knowledge and
559 objective calculations for landslide susceptibility assessment, assuming the use of specific
560 statistical tests if only a few landslide data are available. The proposed procedure has to be
561 tested in other types of environment in order to verify its spatial robustness.
562

563 **Acknowledgements**

564 This research was financially supported by the European Union through the research
565 programme ALARM (Assessment of Landslide Risk and Mitigation in Mountain Areas),
566 contract EVG1-2001-00018, 2002-2004, Coordinator: S. Silvano (CNR-IRPI, Padova).
567

568 **References**

569 Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., Wright, D.F., 1993. Weights of evidence modeling and weighted logistic
570 regression for mineral potential mapping. In: Davis, J.C., Herzfeld, U.C. (Eds.), Computer in Geology, 25 Years of
571 Progress. Oxford University Press, Oxford, pp. 13-32.

572 Agterberg, F.P., Cheng, Q., 2002. Conditional independence test for weights of evidence modeling. Natural Resources
573 Research 11, 249-255.

574 Aleotti, P., Chowdhury, R., 1999. Landslide hazard assessment: summary review and new perspectives. Bulletin of
575 Engineering Geology and the Environment 58, 21-44.

576 Atkinson, P.M., Massari, R., 1998. Generalised linear modelling of susceptibility to landsliding in the central Apennines,
577 Italy. Computers and Geosciences 24, 373-385.

578 Ardizzone, F., Cardinali, M., Carrara, A., Guzzetti, F., Reichenbach, P., 2002. Uncertainty and errors in landslide mapping
579 and landslide hazard assessment. Natural Hazard and Earth System Science 2, 3-14.

580 Begueria, S., Lorente, A., 1999. Landslide Hazard Mapping by Multivariate Statistics: a Comparison of Methods and Case
581 Study in the Spanish Pyrenees. The Damocles Project Work, Contract N° EVG1-CT 1999-00007. Technical Report. 20 pp.

582 Bégueria, S., 2006. Validation and evaluation of predictive models in hazard assessment and risk management. Natural
583 Hazards 17, 315-329.

584 Bonham-Carter, G.F., 1994. Geographic Information System for Geoscientists: Modelling with GIS. Pergamon Press,
585 Oxford. 398 pp.

586 Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1989. Weights of evidence modelling: a new approach to mapping
587 mineral potential. In: Agterberg, F.P., Bonham-Carter, G.F. (Eds.), Statistical Applications in Earth Sciences. Geological
588 Survey of Canada, Ottawa, pp. 171-183.

589 Bonham-Carter, G.F., Agterberg, F.P., Wright, D.F., 1990. Statistical pattern integration for mineral exploration. In: Gaal, G.,
590 Merriam D.F. (Eds.), Computer Applications in Resource Estimation: Prediction and Assessment for Metals and
591 Petroleum. Pergamon Press, Oxford, pp. 1-21.

592 Brabb, E.E., 1984. Innovative approaches to landslide hazard mapping. Proceedings of Fourth International. Symposium on
593 Landslides, Toronto, pp. 307-324.

594 BRGM, 1974. Carte et notice géologique de Barcelonnette au 1:50 000ème, XXXV-39. Bureau des Recherches Géologiques
595 et Minières. Orléans.

596 Carrara, A., Cardinali, M., Guzzetti, F., Reichenbach, P., 1995. GIS technology in mapping landslide hazard. In: Carrara, A.,
597 Guzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer, Dordrecht, pp. 135-176.

598 Chàcon, J., Irigaray, C., Fernàndez, T., El Hamdouni, R., 2006. Engineering geology maps: landslides and geographical
599 information systems. Bulletin of Engineering Geology and the Environment 65, 341-411.

600 Chung, C.F., 2006. likelihood ratio functions for modeling the conditional probability of occurrence of future landslides for
601 risk assessment. Computers and Geosciences, 32, 1052-1068.

602 Chung, C. F., Fabbri, A. G., 2003. Validation of spatial prediction models for landslide hazard mapping. Natural Hazards, 30,
603 451–472.

604 Chung, C.F., Fabbri, A.G., van Westen, C.J., 1995. Multivariate regression analysis for landslide hazard zonation. In:
605 Carrara, A., Guzetti, F. (Eds.), Geographical Information Systems in Assessing Natural Hazards. Kluwer, Dordrecht, pp.
606 107-133.

607 Clerici, A., Perego, S., Tellini, C., Vescovi, P., 2002. A procedure for landslide susceptibility zonation by the conditional
608 analysis method. Geomorphology 48, 349-364.

609 Crozier, M.J., Glade, T., 2005. Landslide hazard and risk: Issues, Concepts and Approach. In: Glade, T., Anderson, M.,
610 Crozier, M.J. (Eds.), Landslide Hazard and Risk. Wiley, Chichester, pp. 1-40.

611 Dai, F.C., Lee, C.F., Ngai, Y.Y., 2002. Landslide risk assessment and management overview. Engineering Geology 64, 65-
612 87.

613 Davis J.C., 2002. Statistics and Data Analysis in Geology, Third Edition. John Wiley & Sons, New York, 638 pp.

614 Dikau, R., Brunsden, D., Schrott, L., Ibsen M-L., 1996. Landslides Recognition, Identification, Movement and Causes. John
615 Wiley & Sons, New York, 251 pp.

616 Dumolard, P., Dubus, N., Charleux, L., 2003. Les statistiques en géographie. Ed. Belin. Paris, 239 pp.

617 Fell, R., 1994. Landslide risk assessment and acceptable risk. Canadian Geotechnical Journal 31, 261-272.

618 Fielding, A.H., Bell, J.F., 1997. A review of methods for the assessment of prediction errors in conservation presence/
619 absence models. Environmental Conservation 24, 38-49.

620 Flageollet, J-C., Maquaire, O., Martin, B., Weber, D., 1999. Landslides and climatic conditions in the Barcelonnette and Vars
621 basins Southern French Alps, France. Geomorphology 30, 65-78.

622 Glade, T., Crozier, M.J., 2005. A review of scale dependency in landslide hazard and risk analysis. In: Glade, T., Anderson,
623 M., Crozier, M.J. (Eds.), Landslide Hazard and Risk. John Wiley and Sons, Chichester, pp. 75-138.

624 Greco, R., Sorriso-Valvo, M., Catalano, E., 2007. Logistic Regression analysis in the evaluation of mass movements
625 susceptibility: the Aspromonte case study, Calabria, Italy. Engineering Geology 89, 47-66.

626 Guzzetti, F., Carrara, A., Cardinali, M., Reichenbach, P., 1999. Landslide hazard evaluation: a review of current techniques
627 and their application in a multi-scale study, central Italy. Geomorphology 31, 181–216.

628 Guzetti, F., Reichenbach, P., Ardizzone, F., Cardinali, M., Galli, M., 2006. Estimating the quality of landslide susceptibility
629 models. Geomorphology 81, 166-184.

630 Hansen, A., 1984. Landslide hazard analysis. In: Brunsden, D., Prior, D.B. (Eds.), Slope Instability. John Wiley & Sons, New
631 York, pp. 532-602.

632 Howell, D.C., 1997. Statistical Methods for Psychology. Fourth Edition. ITP, De Boeck University, 768 pp.

633 Kojima, H., Chung, C.F., Van Westen, C.J., 2000. Strategy on landslide type analysis based on expert knowledge and the
634 quantitative prediction model. International Archives of Photogrammetry & Remote Sensing 33, 701-708.

635 Kemp, L.D., Bonham-Carter, G.F., Raines, G.L., Looney, C.G., 2001. Arc-SDM: Arcview extension for spatial data
636 modelling using weights of evidence, logistic regression, fuzzy logic and neural network analysis,
637 http://ntserv.gis.nrcan.gc.ca/sdm/

638 Kendall, M., Stuart, A., 1979. The Advanced Theory of Statistics: Inference and Relationship. Griffin, London, 748 pp.

639 Leroi, E., 2005. Global rockfalls risk management process in 'La Désirade' Island (French West Indies). Landslides 2, 358-
640 365.

641 Malet, J-P., Van Asch, Th.W.J., van Beek, R., Maquaire, O., 2005. Forecasting the behaviour of complex landslides with a
642 spatially distributed hydrological model. Natural Hazards and Earth System Sciences 5, 71-85.

643 Maquaire, O., Malet, J.P., Remaître, A., Locat, J., Klotz, S., Guillon, J., 2003. Instability conditions of marly hillslopes:
644 towards landsliding or gullying? The case of the Barcelonnette basin, South-East France. Engineering Geology 70, 109-
645 130.

646 MATE/MATL, 1999. Plan de Prévention des Risques (PPR): Risques de Mouvements de terrain, Ministère de
647 l'Aménagement du Territoire et de l'Environnement (MATE), Ministère de l'Equipement des Transports et du Logement
648 (METL). La Documentation Française, Paris. 74 pp.

649 Pistocchi A., Luzi, L., Napolitano, P., 2002. The use of predictive modelling techniques for optimal exploitation of spatial
650 databases: a case study in landslide hazard mapping with expert system-like methods. Environmental Geology 41, 765-
651 775.

652 Remondo, J., Gonzàlez-Diez, A., Dìaz de Teràn, J.R., Cendrero, A., 2003. Landslides susceptibility models utilising spatial
653 data analysis techniques. A case study from the lower Deba Valley, Guipùzcoa (Spain). Natural Hazards 30, 267-279.

654 Soeters R., Van Westen C.J., 1996. Slope instability, recognition, analysis, and zonation. In: Turner, A.K., Schuster, R.L.
655 (Eds.), Landslides Investigation and Mitigation, Transportation Research Board, Special Report 247. National Research
656 Council, Washington, pp. 129-177.

657 Soriso Valvo, M., 2002. Landslides; from inventory to risk. In Rybář, J., Stemberk, J., Wagner, P. (Eds.), Landslides,
658 Proceedings of the International European Conference on Landslides. Balkema, Rotterdam, pp. 79-93.

659 Spiegelhater, D., Kill-Jones, R.P., 1984. Statistical and knowledge approaches to clinical decision-support systems, with an
660 application in gastroenterology. Journal of the Royal Statistical Society 147, 35-77.

661 Sterlacchini S., Maseti M., Poli, S. 2004. Spatial integration of thematic data for predictive landslide mapping: a case study
662 from Oltrepo Pavese area, Italy. In: Lacerda W.A., Ehrlich M., Fountoura, S.A.B., Sayão, A.S.F. (Eds.), Landslides
663 Evaluation and Stabilization. Balkema, Rotterdam, pp. 109-116.

664 Süzen, M.L., Doyuran, V., 2004. Data driven bivariate landslide susceptibility assessment using geographical information
665 systems: method and application to Asarsuyu catchment, Turkey. Engineering Geology 71, 303-321.

666 Thiart, C., Bonham-Carter, G.F., Agterberg, F.P., 2003. Conditional independence in weights of evidence: application of an
667 improved test. IAMG, International Association for Mathematical Geology, September 7th-12th, 2003, Portsmouth,
668 United-Kingdom.

669 Thiery, Y., Puissant, A., Malet, J-P., Remaitre, A., Beck, E., Sterlacchini, S., Maquaire, O., 2003. Towards the construction
670 of a spatial database to manage landslides with GIS in mountainous environment. In: Proceedings of AGILE 2003: The
671 Science behind the Infrastructure, 6th AGILE Conference on Geographic Information Science. 24th-26th, April 2003,
672 Lyon, France, pp. 37-44.

673 Thiery Y., Sterlacchini S., Malet J-P., Puissant A., Maquaire O., 2004. Strategy to reduce subjectivity in landslide
674 susceptibility zonation by GIS in complex mountainous environments. In: Toppen, F., Prastacos, P. (Eds.), Proceedings of
675 AGILE 2004: 7th AGILE Conference on Geographic Information Science. 29th April – 1st May 2004, Heraklion, Greece,
676 pp. 623-634.

677 Thiery, Y., Malet, J-P., Sterlacchini, S., Puissant, A., Maquaire, O., 2005 Analyse spatiale de la susceptibilité des versants
678 aux mouvements de terrain, comparaison de deux approches spatialisées par SIG. Revue internationale de
679 géomatique/European journal of GIS and spatial analysis 15, 227-245.

680 Van den Eeckhaut, M., Vanwalleghem, T., Poesen, J., Govers, G., Verstraeten G., Vandekerckhove, L., 2006. Prediction of
681 landslide susceptibility using rare events logistic regression: a case-study in the Flemish Ardennes (Belgium).
682 Geomorphology 76, 392-410.

683 Van Westen, C.J., 1993. Application of Geographic Information Systems to Landslide Hazard Zonation. ITC Publication,
684 vol. 15. International Institute for Aerospace and Earth Resources Survey, Enschede, 245 pp.

685 Van Westen, C.J., 2004. Geo-Information tools for landslide risk assessment: an overview of recent developments. In
686 Lacerda W.A., Ehrlich M., Fountoura, S.A.B., Sayão, A.S.F. (Eds.), Landslides Evaluation and Stabilization. Balkema,
687 Rotterdam, pp. 39-56.

688 Van Westen C.J., Rengers N., Soeters R., 2003. Use of geomorphological information in indirect landslide susceptibility
689 assessment. Natural Hazards 30, 399-419.

690 Van Westen C.J., Van Asch, Th.W.J., Soeters, R., 2006. Landslide hazard and risk zonation: why is it still so difficult?
691 Bulletin of Engineering Geology and the Environment 65, 167-184.

692 Varnes, D.J., 1984, Landslide Hazard Zonation, a Review of Principles and Practice. IAEG Commission on Landslides,
693 UNESCO, Paris. 63 pp.

694 Wills C.J., McCrinck, T.P., 2002. Comparing landslide inventories: the map depends on the method. Environmental and
695 Engineering Geoscience 8, 279-293.

696 Yin, K.L., Yan, T.Z., 1988. Statistical prediction model for slope instability of metamorphosed rocks. In: Bonnard, C. (Ed.),
697 Landslides, Proceedings of Fifth International Symposium in Landslides. Balkema, Rotterdam, pp. 1269-1272.

Tables

Table 1. Morphometric characteristics of active landslides inventoried with a high mapping confidence index (MCI). μ is geometric average; σ is standard deviation.

Landslide type	Number	Depth (m)		Width (m)		Length (m)		Slope of LTZ (°)		Landslide size (m^2)		Size of LTZ (m^2)	
		μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Shallow translational slide	50	2	0.6	60	25	77	70	31	9	2766	2389	866	714
Rotational slide	54	6	3	140	136	118	114	21	9	12527	12971	4601	3947
Translational slide	88	6.5	4.5	78	70	217	168	21	6	14874	19002	4400	4100

Table 2. Thematic data used in the statistical model.

Themes	Map	Source of information and methods used
Landslide inventory	Landslide inventory map (LAI)	API (air-photo interpretation), field survey, analysis of available documents
Relief	Slope gradient map (SLO)	DTM elaborated by digitization and interpolation of elevation lines extracted from topographical maps (1:10,000)
	Slope curvature map (CUR)	
Geology	Lithological map (LIT)	Analysis of geological map, field survey
	Surficial formation map (SF)	Analysis of geological and geomorphological maps, field survey
	Thickness map (TSF)	Field survey
	Bedding map (BED)	Analysis of geological map, field survey
Hydrology	Distance to stream map (HYD)	API, analysis of topographical maps
Landuse	Landuse map (LAD)	SIA (satellite imagery analysis), API, field survey

Table 3. Expert rules and associated environmental conditions used to define the direct susceptibility map. SLO: slope gradient; LAD: land use; CUR: slope curvature.

Susceptibility class	Expert rule	Environmental conditions
S0: no susceptibility	Environmental conditions favourable to slope stability. No possibility of landslide developments for the next one hundred years.	SLO: 0-10° LAD: arable land, permanent crop
S1: low susceptibility	Environmental conditions are slightly favourable to slope instability. Low possibility of landslide developments for the next one hundred years. Future human and socio-economic developments of the area are possible and subject to specific attention.	SLO: 10-20° LAD: pasture, grassland CUR: moderate presence of topographic irregularities
S2: moderate susceptibility	Environmental conditions are moderately favourable to slope instability. Moderate possibilities of landslide developments for the next one hundred years. Mitigation works are essential for future human and socio-ecomonic developments of the area.	SLO: 20-30° LAD: pasture, grassland, forests lowly maintained CUR: high presence of topographic irregularities , hummocky topography
S3: high susceptibility	Environmental conditions are very favourable to slope instability. High possibility of landslide developments for the next one hundred years. Future human and socio-ecomonic developments of the area are impossible.	SLO: > 30° LAD: landuse highly deteriorated, bare soils, forests not maintained CUR: very hummocky topography

Table 4. Characteristics of the response variable (RV) introduced in the statistical model to identify the most relevant spatial locations of cells to represent the variability of predisposing factors within LTZs, and relative error associated with the simulations. The simulations were performed with the predictive variables (PVs) SLO, SF, LIT and LAND (Table 2). STS: shallow translational slides; RS: rotational slides; TS: translational slides.

Response variable (RV)	Characteristics of the response variable	Relative error ξ (–)		
		STS	RS	TS
RV-1	Use of all (460) cells of the landslide triggering zones (LTZs)	0.50	0.54	0.45
RV-2	Use of the centre of mass of each LTZ (e.g. one cell per LTZ)	0.76	0.73	0.74
RV-3	Use of the total number of cells in a radius of 10 m around RV-2 (e.g. 230 cells)	0.57	0.60	0.49
RV-4	Use of the total number of cells of small LTZs (mean size: TS: 215 m^2; RS: 260 m^2; STS: 60 m^2)	0.64	0.69	0.69
RV-5	Use of the total number of cells of medium-size LTZs (mean size: TS: 400 m^2; RS: 450 m^2; STS: 65 m^2)	0.58	0.62	0.52
RV-6	Use of the total number of cells of large LTZs (mean size: TS: 650 m^2; RS: 640 m^2; STS: 190 m^2)	0.53	0.54	0.46
RV-7	Use of the cells representing the most frequent combination of PVs observed in each LTZ (e.g. 230 cells)	0.45	0.43	0.40

Table 5. Confusion matrix. a: true positives; b: false positives; c: false negatives; d: true negatives.

		Observed	
		X_1	X_0
Predicted	X'_1	A	b
	X'_0	C	d

Table 6. Statistics derived from the confusion matrix. N: number of cells in the study area. a: true positives; b: false positives; c: false negatives; d: true negatives.

Correct classification rate	(a + d) / N	Proportion of correctly classified observations
Misclassification rate	(b + c) / N	Proportion of incorrectly classified observations
Sensitivity	a / (a + c)	Proportion of positive cases correctly predicted
Specificity	d / (b + d)	Proportion of negative cases correctly predicted
Kappa (*K*) coefficient	[(a + d) - (((a + c)(a + b) + (b + d)(c + d)) / N)] / [N - (((a + c)(a + b)+(b + d)(c + d)) / N]	Proportion of specific agreement

Table 7. Example of association measures between RV-7 and PVs for the translational slides. The PVs CUR, HYD and BED are not introduced in the model because there is no causal relation between the occurrence of the translational slides and these PVs. The bold font indicates the PV used to build an nPV. χ^2-test: from left to right, calculated χ^2, theoretical χ^2, and degree of freedom. The grey-coloured box represents the conditional dependence between PVs and the null hypothesis H_0 rejected for a level of significance $\alpha = 0.05$. Cramer's V coefficient: the bold font indicates moderate to high association between the variables.

PV		LIT	**SF**	TSF	LAD	CUR
SLO	χ^2	2.6 *12.5* (6)	33.1 *21* (12)	104.3 *28.8* (18)	75.5 *36.4* (24)	81.6 *21* (12)
	V	0.11	**0.42**	0.26	0.29	**0.41**
LIT	χ_2	-	0.2 *5.9* (2)	5.7 *7.8* (3)	0.2 *9.5* (4)	1.2 *5.9* (2)
	V	-	0	0.15	0.03	0.07
SF	χ^2	-	-	9.6 *12.5* (6)	35.3 *15.5* (8)	7.2 *9.4* (6)
	V	-	-	0.14	0.27	0.12
TSF	χ^2	-	-	-	31.8 *21* (12)	55.7 *12.6* (6)
	V	-	-	-	0.2	0.38
LAD	χ^2	-	-	-	-	24.5 *9.5* (4)
	V	-	-	-	-	0.23

Table 8. Contribution of PVs on the explained variance of the axes F1 to F4 for three landslide types. The most contributive PVs for each axis are indicated in grey and are used to define nPVs. The classes chosen to build nPVs are detailed in the last column.

	SLO	LIT	SF	TSF	LAD	CUR	HYD	BED	Explained variance (%)	Structure of nPVs
Shallow translational slides										
F1	25.6	18.6	19.1	3.2	15.3	0.1	0.1	18.2	13.1	
F2	10.3	6.9	5.9	21.5	25.9	3.6	0.2	25.6	22.7	nPV-1: SLO (15-25°, 25-35°, 35-45°, 45-55°) + SF (colluvium, scree, moraine deposit)
F3	18.4	16.5	6.9	25.0	16.4	2.0	10.4	4.5	32	
F4	21.64	3.3	10.0	28.1	28.5	7.4	0.7	0.3	40.5	
Rotational slides										
F1	33.1	17.2	24.9	3.0	21.0	0.6	0.05	-	16.4	
F2	24.9	3.8	0.8	34.2	7.8	19.7	2.4	-	28.4	nPV-3: SLO (10-20°, 20-30°, 30-40°) + SF (all classes)
F3	19.8	17.0	6.7	29.7	9.2	11.2	6.4	-	39.8	
F4	40.9	0.9	4.5	22.5	20.6	0.7	10.0	-	49.3	
Translational slides										
F1	37.1	0.4	25.6	21.1	15.7	-	-	-	12.9	nPV-3: SLO (5-15°, 15-25°, 25-35°, 35-45°) + SF (moraine deposit)
F2	36.8	3.1	6.3	29.9	20.6	-	-	-	25.3	
F3	39.9	0.1	12.2	33.8	13.7	-	-	-	36.1	nPV-4: SLO (25-35°, 35-45°) + SF (colluvium or weathered marl)
F4	24.3	2.4	25.7	16.7	30.8	-	-	-	46.0	

Table 9. Relative error ξ and CI results for the best combination of PVs and nPVs

Landslide type	Combination	ξ	χ^2-*test*	*V-coefficient*
Shallow translational slides (STS)	nPV-1 + LAD	0.40	Yes	Low
	nPV-1 + LAD + HYD	0.35	Yes	Low
	nPV-1 + LAD + HYD + CUR	0.21	Yes	Low
	nPV-1 + LAD + HYD + CUR + BED	0.14	Yes	Low
Rotational slides (RS)	nPV-2 + LAD	0.21	Yes	Low
	nPV-2 + LAD + HYD	0.18	Yes	Low
	nPV-2 + LAD + HYD + CUR	0.16	Yes	Low
Translational slides (TS)	nPV-3 + LIT	0.35	Yes	Low
	nPV-3 + LIT + LAD	0.18	Yes	Low

Table 10. Relative error ξ of the best simulations for the 'sampling area' and the total study area. Results are indicated for the LTZ and the total area of landslide (L). Simulations are computed with RV-7.

	STS (nPV-1 + LAD + HYD + CUR + BED)		RS (nPV-2 + LAD + HYD + CUR)		TS (nPV-3 + LIT + LAD)	
	LTZ	L	LTZ	L	LTZ	L
ξ : 'sampling area'	0.14	0.09	0.16	0.34	0.18	0.41
ξ : total study area	0.22	0.26	0.21	0.33	0.23	0.47

Table 11. Statistical accuracy tests between the indirect and direct susceptibility maps. ccr is the correct classification rate; mcr is misclassification rate.

	Susceptibility class				
	Null	Low	Moderate	High	Global
ccr	0.73	0.81	0.85	0.91	0.61
mcr	0.27	0.19	0.15	0.09	0.39
sensitivity	0.87	0.18	0.08	0.80	0.61
specificity	0.39	0.89	0.95	0.93	0.89
Kappa K	0.36	0.08	0.03	0.43	0.41

Fig. 1. Shaded relief map of the north-facing hillslope of the Barcelonnette Basin and distribution of landslides.

Fig. 2. Landuse map of the north-facing hillslope of the Barcelonnette Basin.

Fig. 3. Simplified geological map (A) and observed landslide types in the Barcelonnette Basin: (B) shallow translational slide in the Abriès Torrent; (C) rotational slide in the Poche Torrent; and (D) translational slide in the Bois Noir catchment.

Fig. 4. Characteristics of the active landslides observed in the Barcelonnette Basin.

Fig. 5. Strategy for susceptibility assessment with the bivariate WOE model at 1:10,000 scale.

Fig. 6. Distribution of landslides and environmental characteristics of the 'sampling area'. (A) inventory of active landslides; (B) slope gradient map; (C) surficial formations map; (D) lithological map; (E) landuse map; (F) thickness of surficial formations map; (G) irregularities of terrain map; (H) outcrop and type of dip map.

Fig. 7. Relative error ξ of the simulations for several quantities of RV cells introduced in the statistical model.

Fig. 8. Cumulative curves of the best simulation obtained in the 'sampling area'. (A) translational slides; (B) rotational slides; (C) shallow translational slides. The susceptibility classes are defined on the basis of the thresholds observed in the cumulative curves of total probabilities. The number of cells in the highest susceptibility class is compared to the distribution of LTZs (relative error ξ).

Fig. 9. Example of WOE simulations for shallow translational slides performed without (A) and with (B) the introduction of an nPV: Statistical simulations with the PVs SLO + SF + LIT + LAD and with the PVs nPV-1 + LAD + CUR + BED, respectively.

Fig. 10. Indirect susceptibility map for the landslide types observed on the north-facing hillslope of the Barcelonnette Basin. (A) shallow translational slides; (B) rotational slides; (C) translational slides.

Fig. 11. Direct susceptibility map produced with the French Official Method of Landslide Risk Zoning.

Fig. 12. Final indirect susceptibility map produced by combining the three indirect landslide susceptibility maps.

Fig. 13. Differences between the direct and final indirect susceptibility maps (example of the high susceptibility class).

Figure 1

Click here to download high resolution image

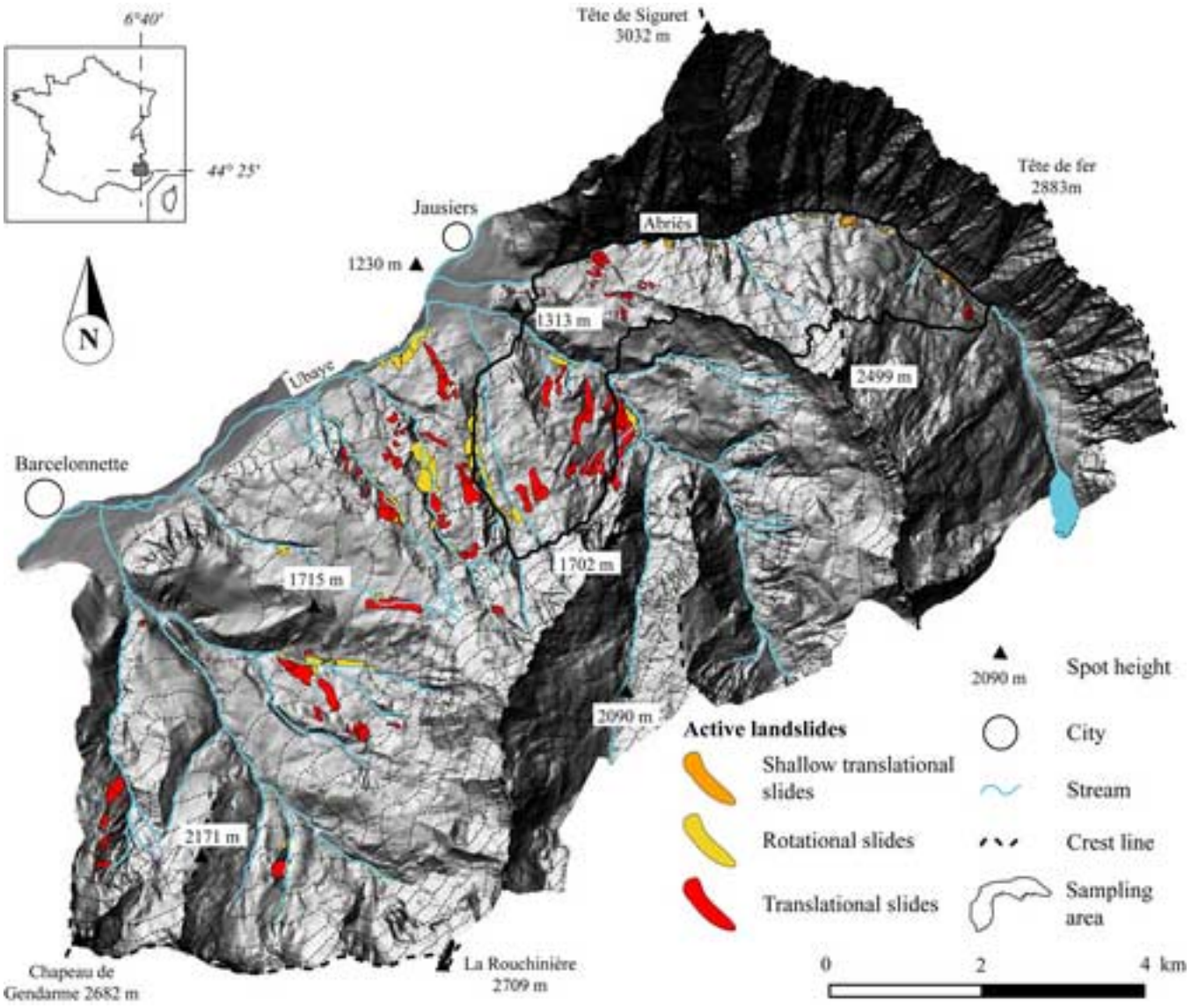

Figure 2

Click here to download high resolution image

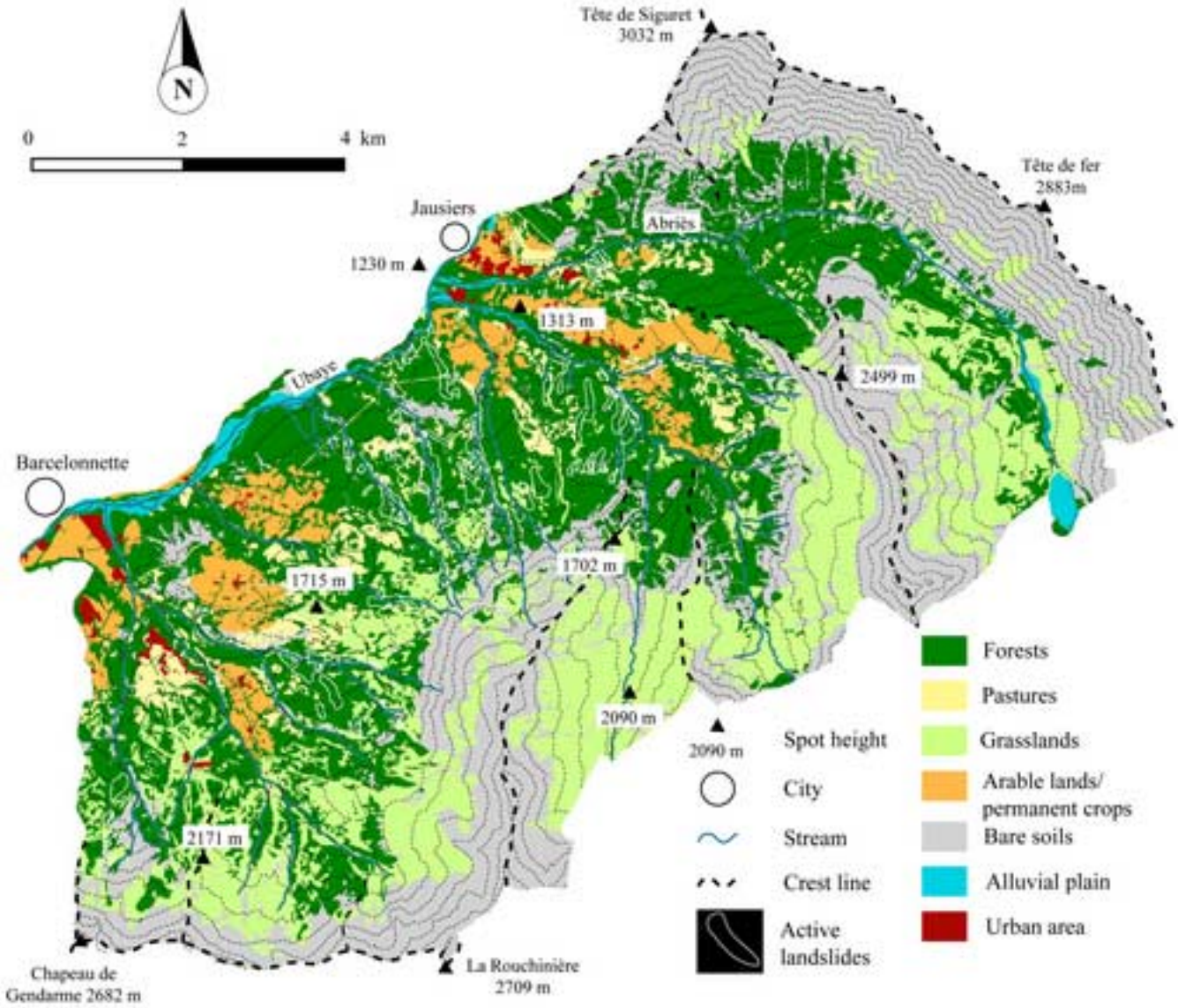

Figure 3

Click here to download high resolution image

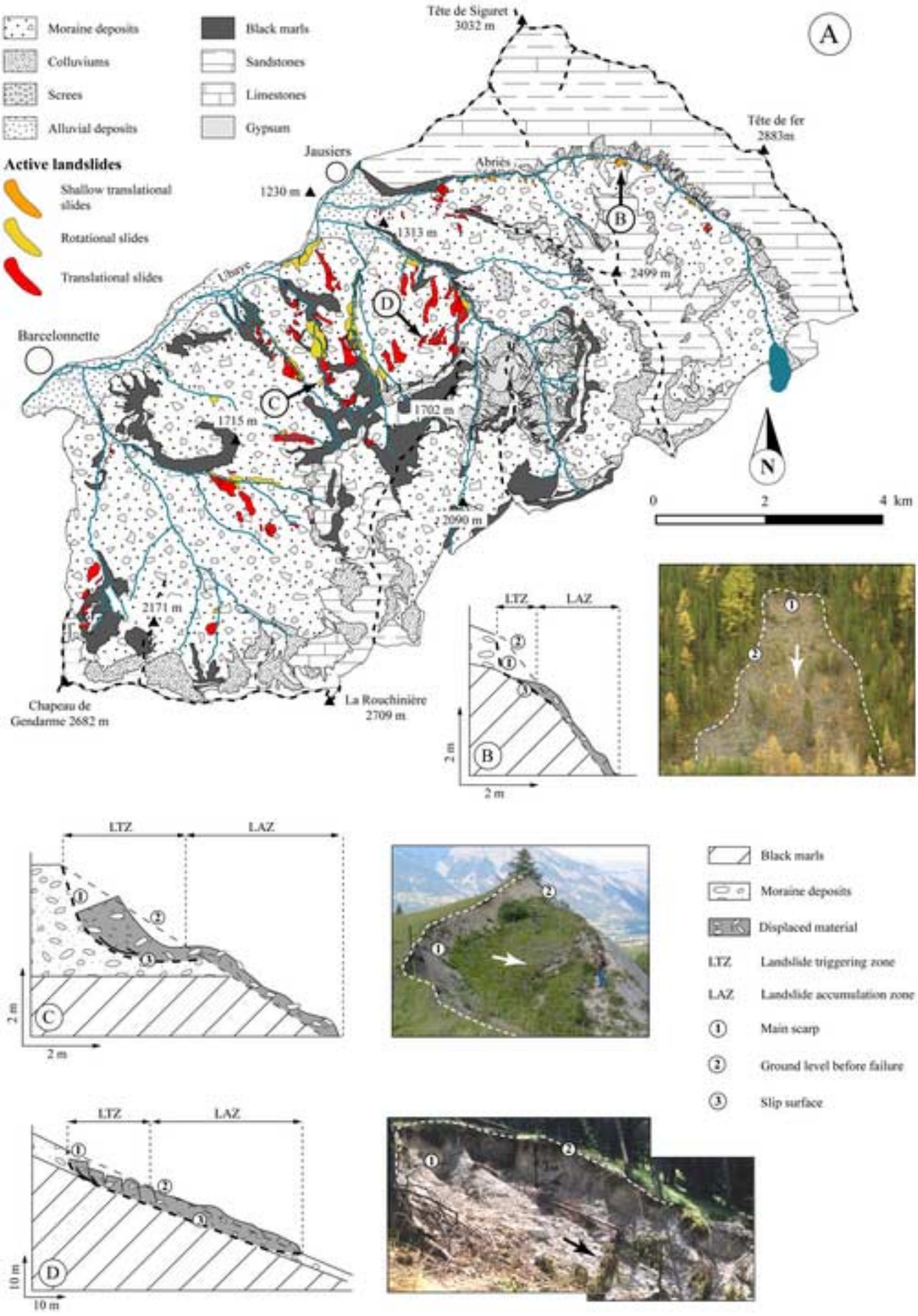

Figure 4

Click here to download high resolution image

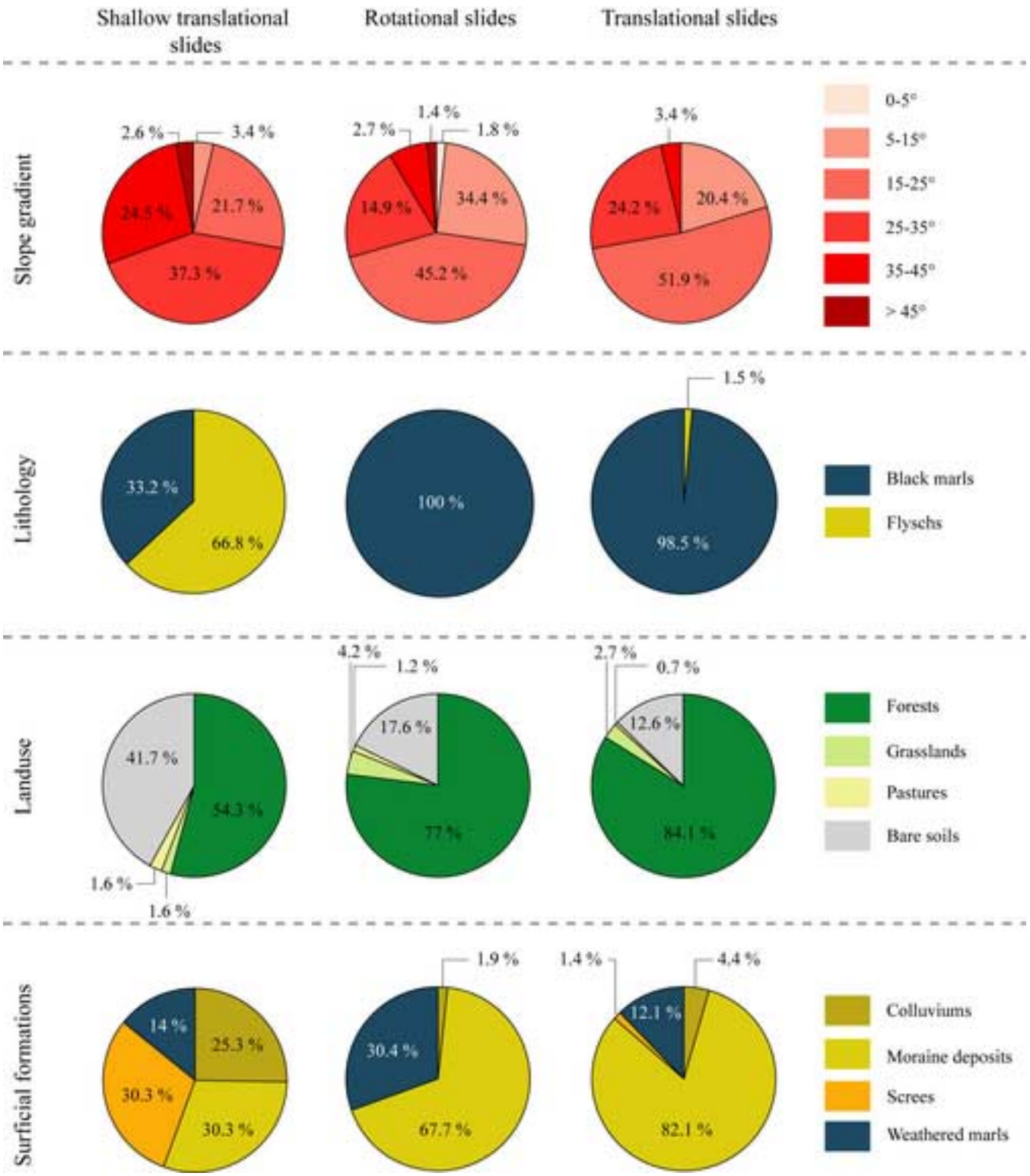

Figure 5

Click here to download high resolution image

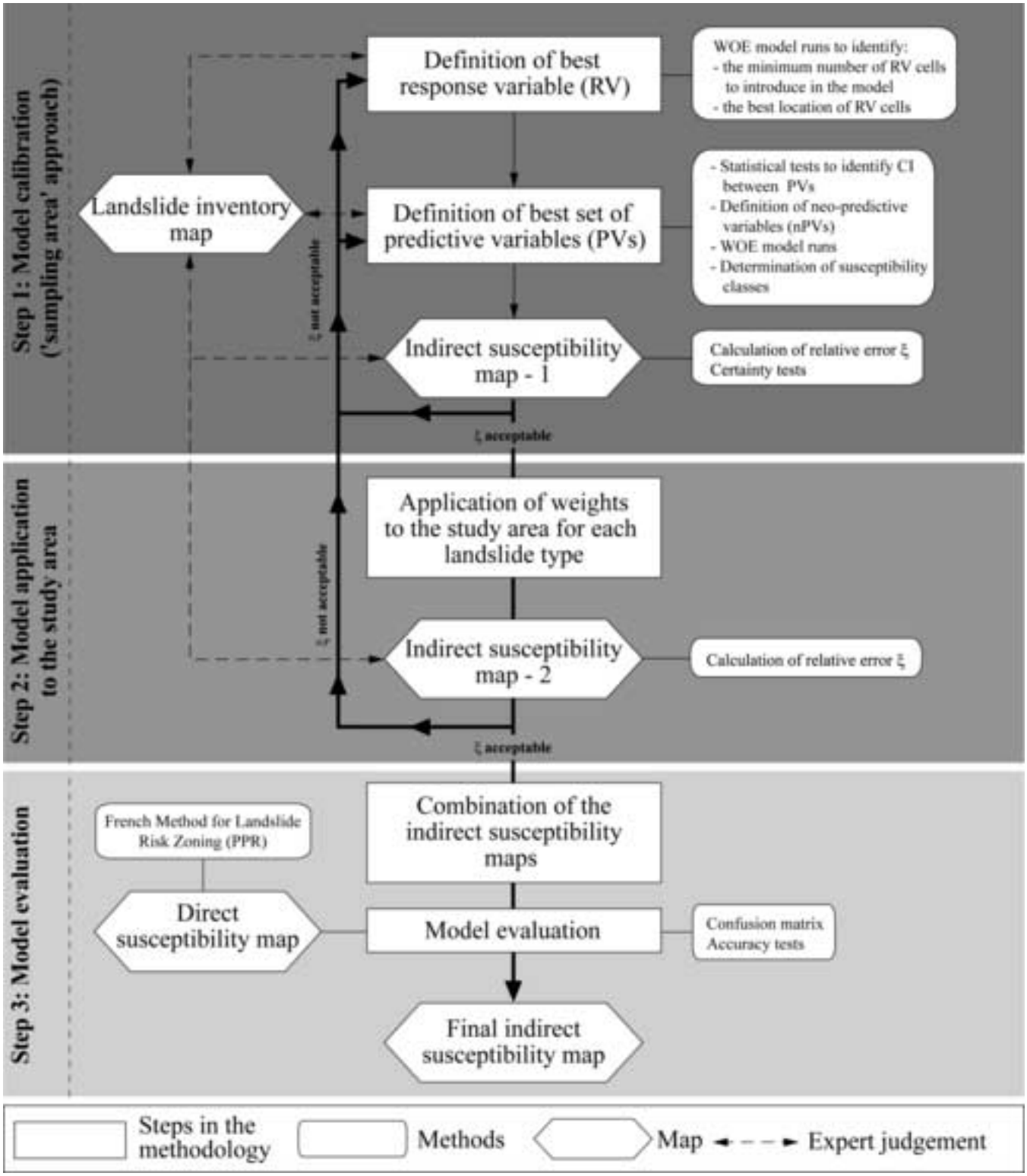

Figure 6

Click here to download high resolution image

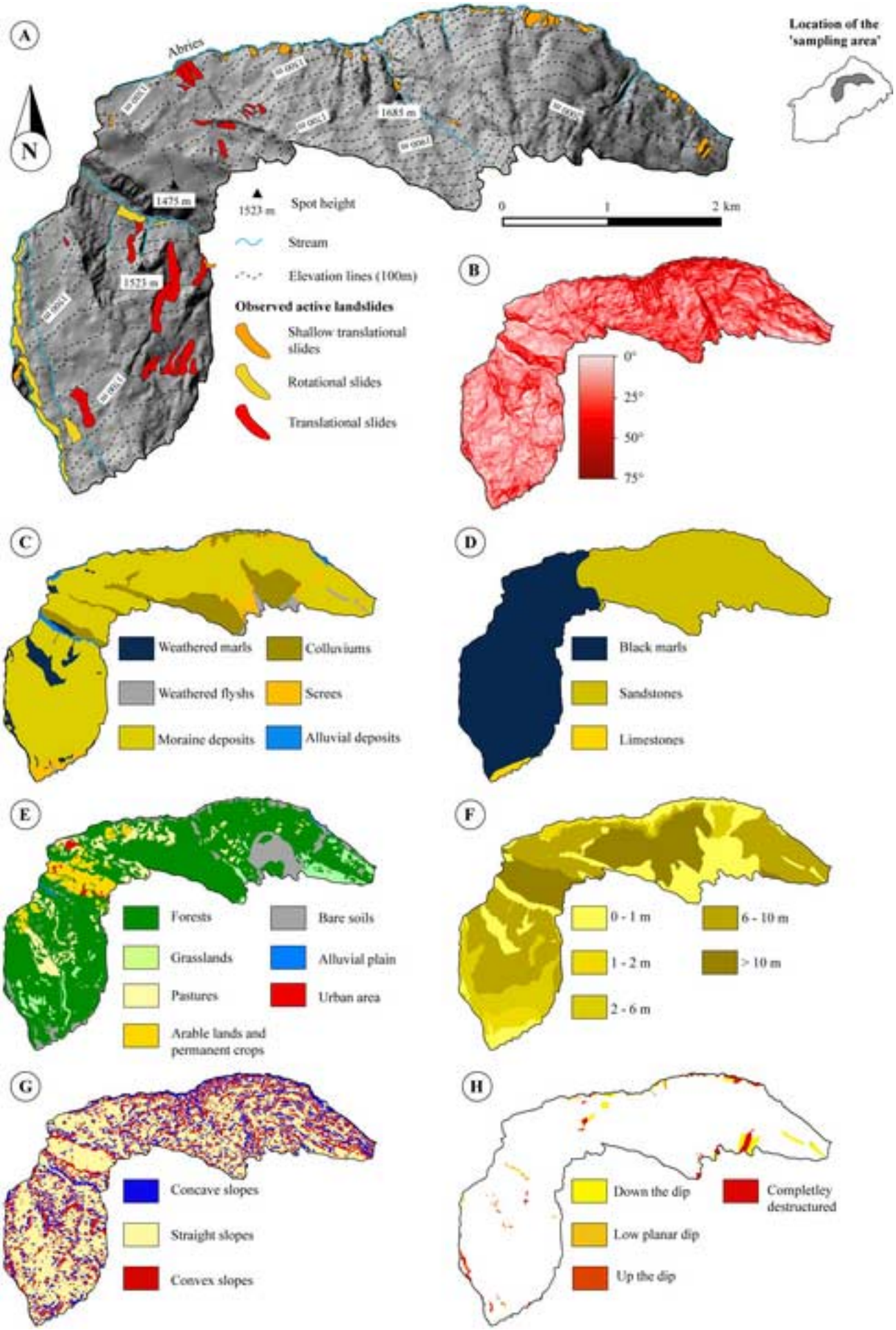

Figure 7
Click here to download high resolution image

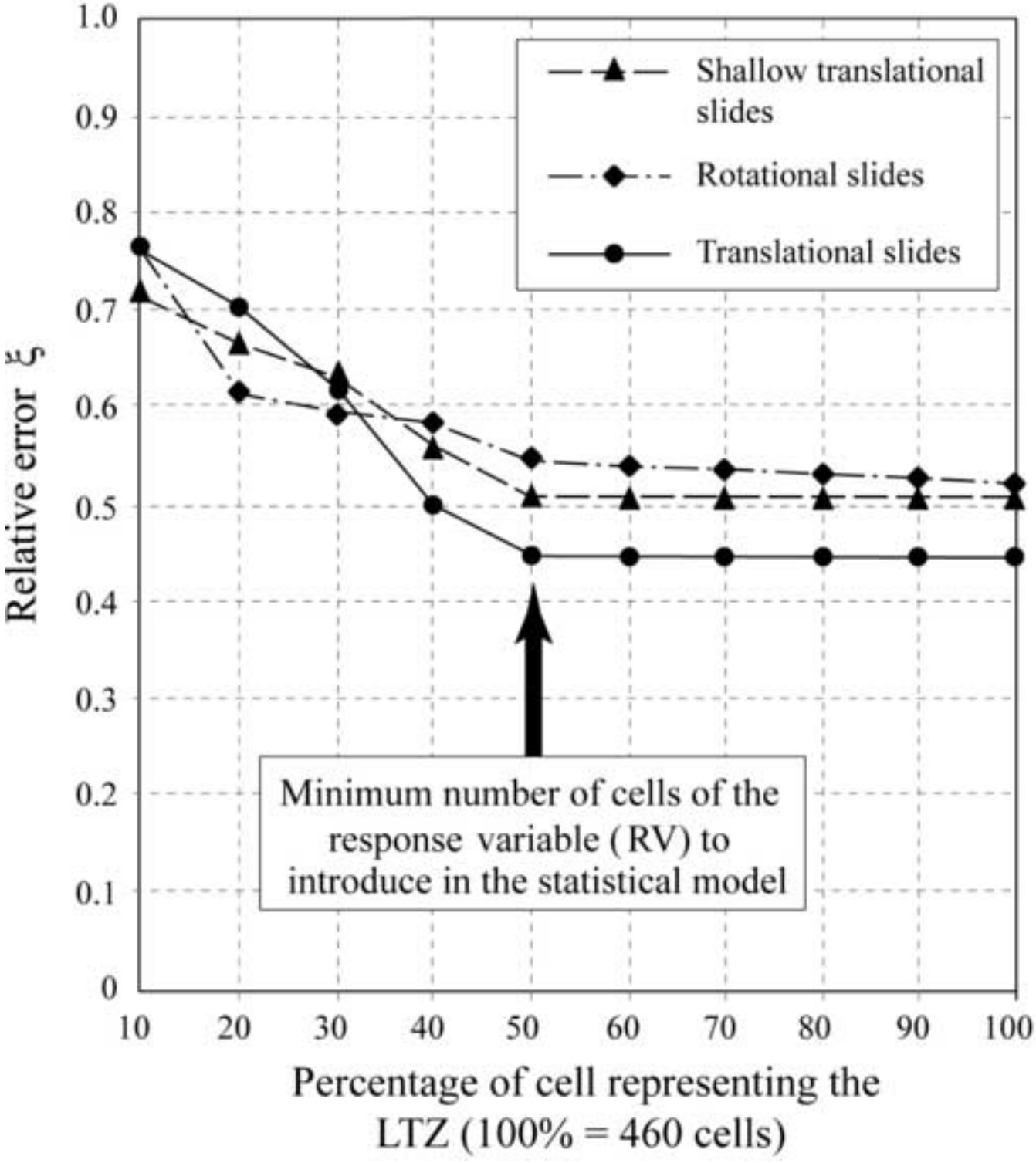

Figure 8

Click here to download high resolution image

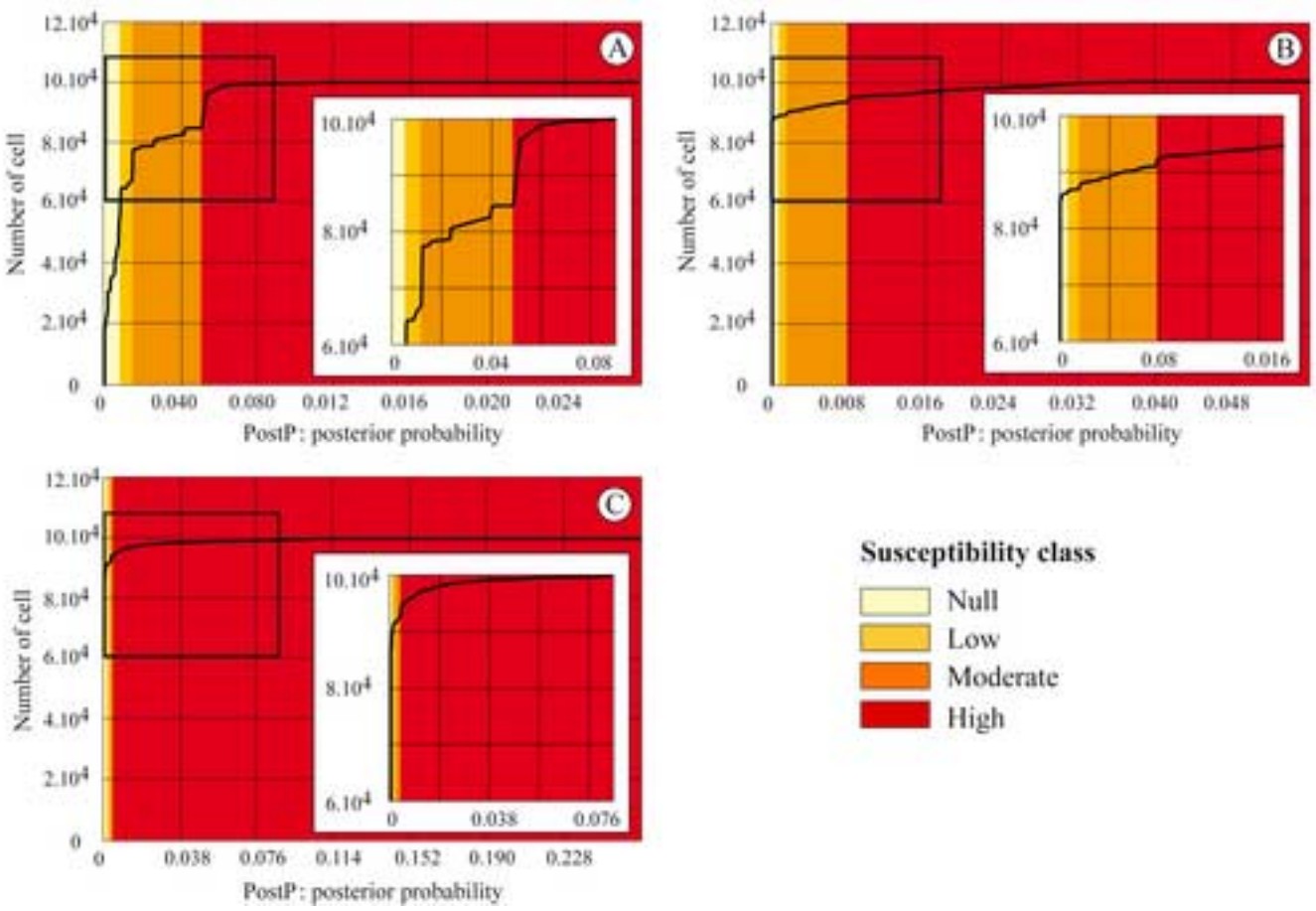

Figure 9

Click here to download high resolution image

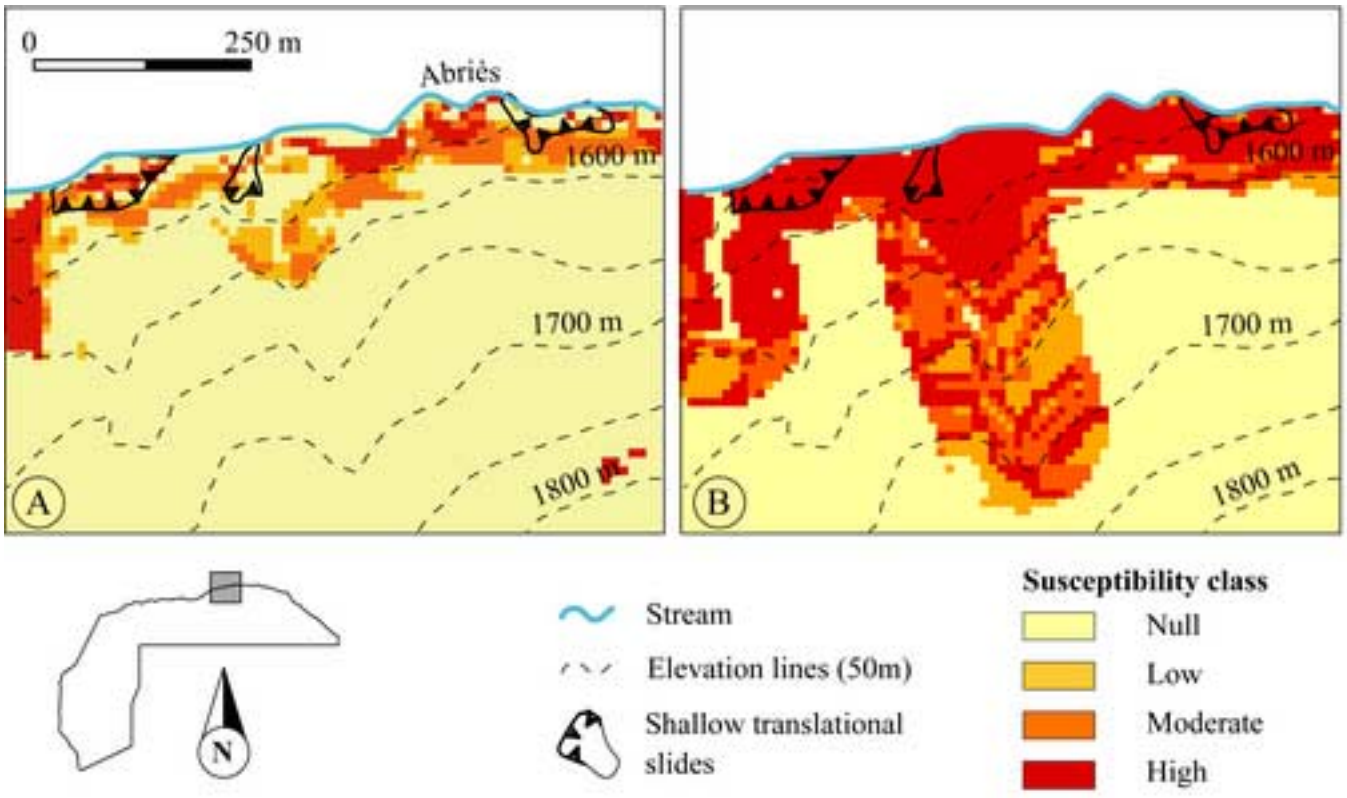

Figure 10

Click here to download high resolution image

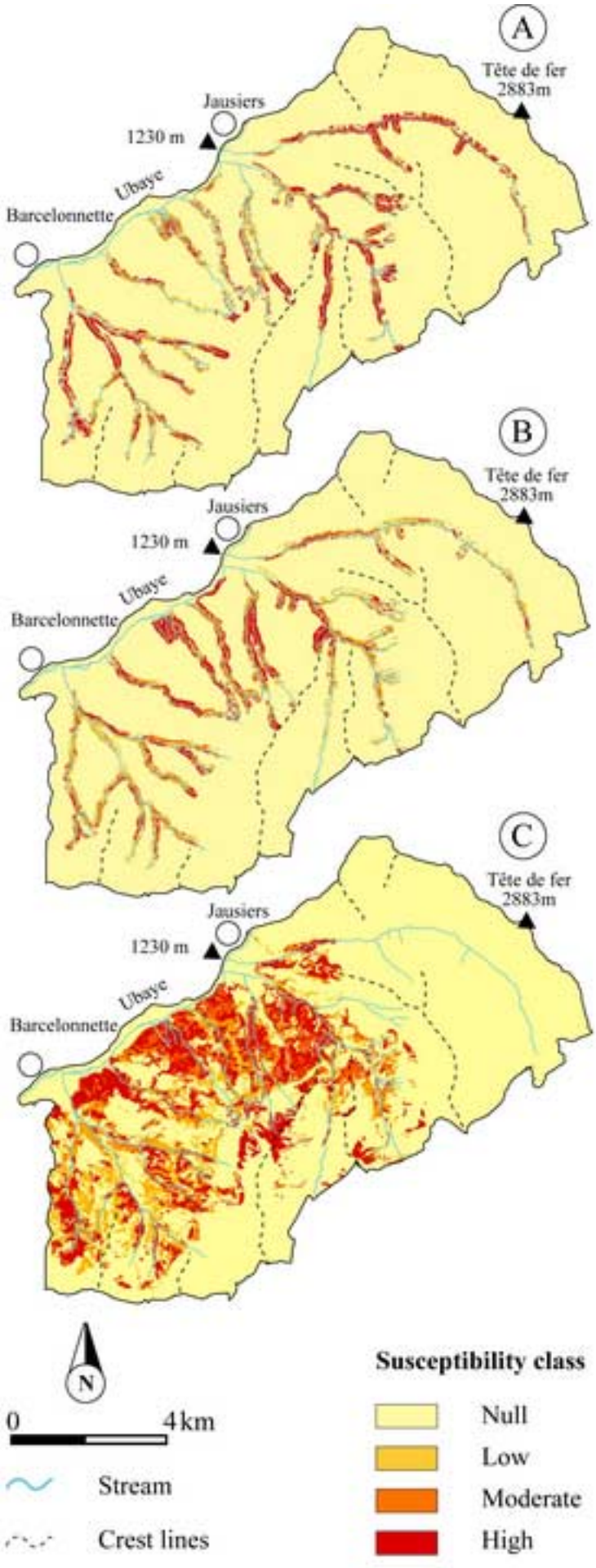

Figure 11

Click here to download high resolution image

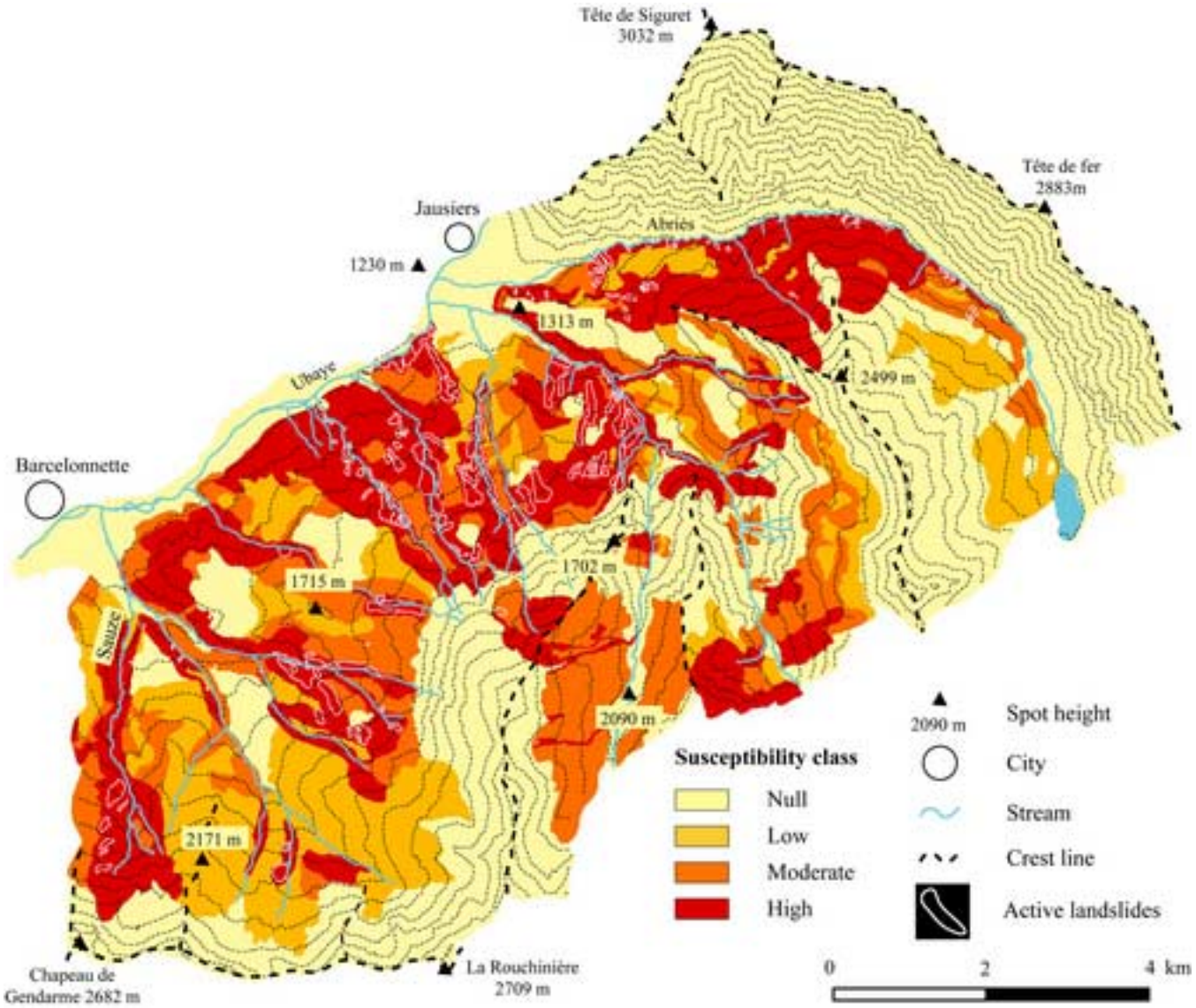

Figure 12

Click here to download high resolution image

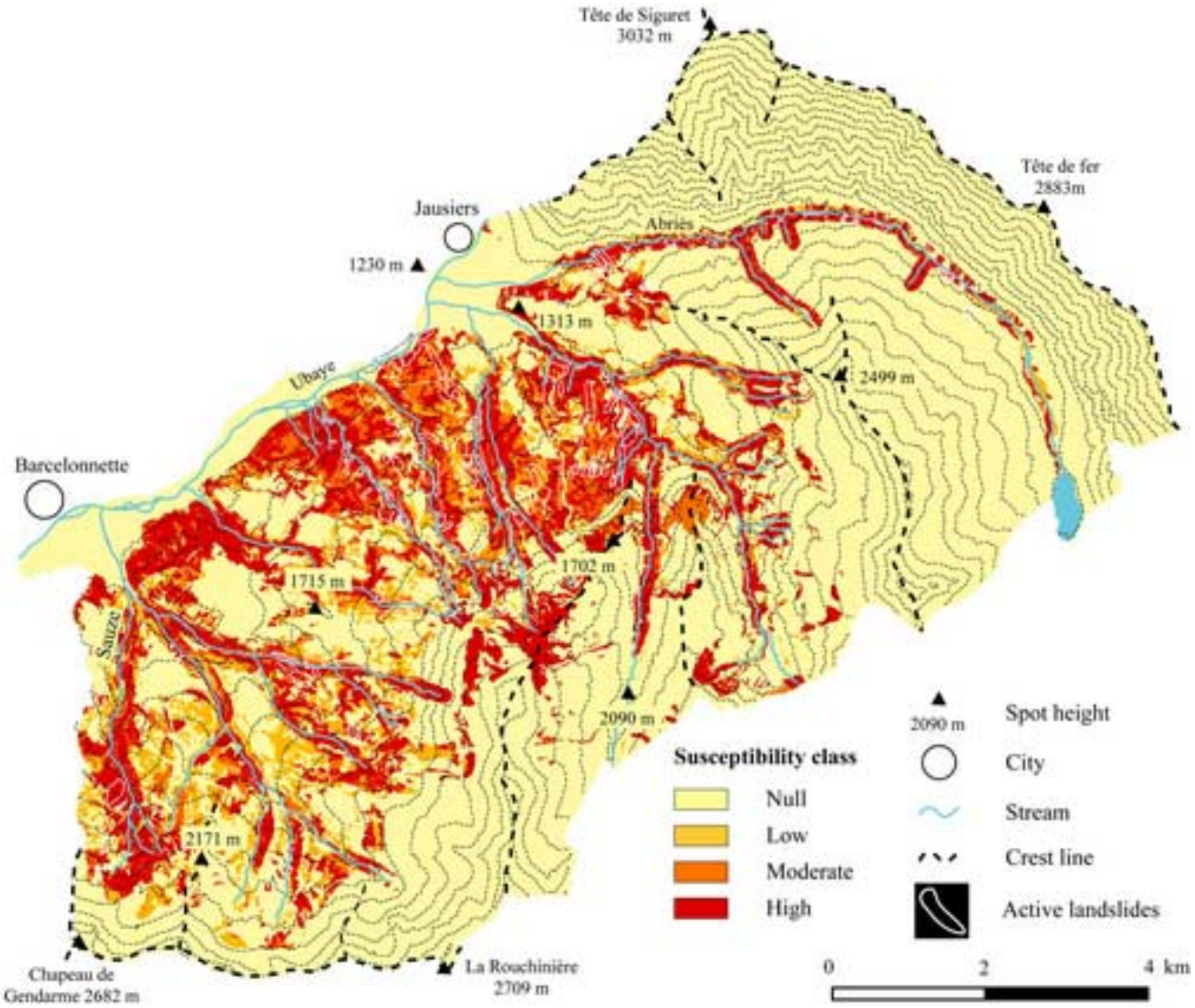

Figure 13

Click here to download high resolution image